Pseudomonas aeruginosa population structure revisited under environmental focus

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades einer
Doktorin der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von Kateryna Selezska
aus Kiev, Ukraine

1. Referent: Professor Dr. Kenneth Nigel Timmis
2. Referent Professor Dr. Dieter Jahn
eingereicht am: 05.07.2010
mündliche Prüfung (Disputation) am: 24.08.2010

Druckjahr 2010

That which does not kill us

makes us stronger

Friedrich Nietzsche

Table of Contents

Summary

Here it is reported a population genetic study of more than 430 environmental strains of *P. aeruginosa*, mostly from river systems in Northern Germany but also from worldwide distributed sampling places. We identify several genetic lineages, of which one is apparently of a predominantly environmental nature. We observe a non-random distribution in environmental habitats and we identified ecological criteria which may be explanatory for this. The results suggest that different selective pressures aid to drive the cladogenic split of evolutionary lineages within *P. aeruginosa*. This study shifts the clinically motivated but from a global perspective rather narrow focus of *P. aeruginosa* as an opportunistic pathogen to its rather true nature as an environmental organism.

1. Introduction

1.1. Short history summary on theory of evolution and evolution driving forces

At the beginning of the 19th century, a French naturalist named Jean Baptiste Pierre Antoine de Monet, Chevalier de Lamarck proposed a full-blown theory of evolution, one of the earliest theories. There, an alchemical complexifying force drove organisms up a ladder of complexity, and a second environmental force adapted them to local environments through use and disuse of characteristics, differentiating them from other organisms (Gould, 2002). Charles Darwin expressed in 1859 in his “On the Origin of Species” a different mechanism for the modus of species evolution. The entire diversity of life forms “from so simple a beginning” was depicted to originate, probably, from the common ancestor (Darwin, 1859). However, Darwin’s ideas lacked the vital basis of genetics, as at this days Mendel’s works on the mechanisms of heredity appeared unnoticed. Only in the first third of 20th century Mendel’s laws were rediscovered and the next stepping stone was put to the establishment of the genetic basis of evolution and hence, population genetics started its development. Theoretical and experimental works in genetics consequently brought new light into the understanding of evolution, and the so-called Modern Synthesis of evolutionary biology arose (Dobzhansky, 1937; Huxley, 1942; Mayr, 1944; Simpson, 1944). Nowadays evolutionary biology faces new challenges and perspectives for the next contemplative breakthrough stepping stone (Rose & Oakley, 2007). The modern stage could be referred to as Evolutionary Biology in the Light of Genomics by the analogy to the Modern Synthesis described as Darwinism in the Light of Genetics (Koonin, 2009).

Evolution is regarded as the change of phenotypes in the course of generations of populations (Mayr, 1997). Microevolution is that part of evolutionary science which deals with changes at the level of species or below, whereas macroevolution deals with the changes that finally lead to the characteristics of taxa higher than species. The current definition and meaning of this term was established in the third decade of the last century with the publication of “Variabilität und Variation” by Yuri Filipchenko (Filipchenko, 1927). Microevolutionary processes may be categorized in two ways: anagenesis and cladogenesis. Anagenesis is regarded as the phyletic change of the structure of a single lineage over time, without the emergence of another, novel, lineage. Cladogenesis, in contrast, is the splitting event of one genetic lineage into two or more genetically different lineages. In this way two or more new forms (species) evolve from one common ancestor.

Both of these processes are governed and driven by the population genetic forces of mutation, recombination, natural selection and genetic drift (Hartl & and Clark, 2007; Lynch, 2007a; Lynch, 2007b)

Natural selection is the process by which an organism that displays favorable characteristics has a higher probability to leave descendants and pass on its genes than other organisms not representing this characteristic. Natural selection acts on the phenotype (Mayr, 1997), but this phenotype needs to be inheritable, as the genetical material is the baggage that descendants get. Thus, the phenotype is being selected and natural selection removes variability from the group.

Mutation is the change in the DNA of the organism. Mutations can be classified as highly deleterious, weakly deleterious, nearly neutral, neutral, weakly advantageous, strongly advantageous (Kimura, 1983). Irrespectively, mutation is a major source of variation on which selection then can act and, thus, lead to adaptation [the process by which a population becomes better suited to its environment (Orr, 2005)]. The critical point for the molecular evolution discussions is the fixation probability of the given mutation. According to the selection theory (neo-Darwinism), most of the occurring mutations are deleterious, and only few are advantageous. Those advantageous mutations are selected and, thus, are being rapidly fixed in the population (i.e., are present in more than 95% of all individuals). Selection here is the only force that contributes to the mutation's fixation in the population. Thus, neo-Darwinism regards the majority of genetic diversity in a population to be the result of adaptive process. The neutral theory (Kimura, 1968) assumes that considerable quota of the mutations are neutral, only few are advantageous, and the rest of the mutations are deleterious. Deleterious mutations are removed from the population, while neutral mutations are fixed in the population by genetic drift, which is the impact of pure chance. According to the theory, genetic drift and selection are the forces that influence mutations' fixation depending on the size of population. Drift mostly influences small populations and is less strong in large populations, whereas selection is weak in small populations but strong in large populations (Hartl & and Clark, 2007). This means that deleterious or slightly deleterious mutations are selected against in big populations, whereas advantageous mutations become fixed. In contrast, in small populations the power of selection is substantially reduced, which enables also slightly deleterious (near-neutral) mutations to become fixed. The Neutral Theory does not neglect the action of selection; however, the Neutral Theory states that the majority of molecular genetic diversity in a population is of neutral or near-neutral character, but not of adaptive value (Ohta, 1973).

Recombination is the exchange of genetic material between organisms (Ochman *et al.*, 2000). Bacteria can exchange genetic material via several mechanisms: natural transformation, conjugation and transduction (Thomas & Nielsen, 2005). Natural transformation is the process of the uptake of free DNA from the environment and its heritable incorporation into the genome by naturally competent bacteria. Conjugation is possible when conjugative plasmids or transposons mediate formation of a pilus which allows a physical contact between the recipient and donor cells. This way the recipient cell receives the donor chromosomal DNA. Transduction is the process which is possible only via interaction of bacteriophage and bacterial cell, when phage while infecting new cell releases DNA into the new host (Claverys *et al.*, 2000; Dubnau, 1999; Matic *et al.*, 1996). Recombination can only be successful if the number of circumstances are sufficed. Those include: availability of the donor DNA to the recipient; the actual process of the uptake of the donor DNA by the recipient cell; escape from the recipient's restriction enzymes; formation of a donor-recipient heteroduplex DNA molecule; escape of the heteroduplex from the host mismatch repair system; and functionality of the donor gene product in the new genetic background (Majewski, 2001). Recombination potential differs dramatically within bacteria: from highly panmictic *Helicobacter pylori* to highly clonal *Buchnera* (Falush *et al.*, 2001; Vos & Didelot, 2009), consequently its influence on the microevolution process differs from population to population. For example, in naturally transformable large populations, a single nucleotide change is around 50 times more seemingly to occur through recombination than through mutation (Feil *et al.*, 2000; Guttman & Dykhuizen, 1994). Thus, recombination (parasexuality) is another major population genetic force that increases variability in prokaryotes. Besides mere increase in variability, when the recombination rate strongly exceeds the mutation rate, recombination has a cohesive effect, which may avoid or at least hinder the cladogenic split of lineages.

The numerous combinations in which these four forces can interact at varying quantities to determine the fate of a population can result in different evolutionary styles according to which organismic lineages may evolve (Cohan & Perry, 2007).

Microevolution of different species has been under the close interest for quite a while, and, thus, specific tendencies in species evolution were recovered for some bacteria. In the discussion about species microevolution and speciation process, it is necessary to clarify the term "species". Thus, before discussing any concrete examples of such bacterial species, it is necessary to take a closer look on what, actually, is a species from the point of view of microbiologists and taxonomists, which concepts and ideas exist and which driving force(s) in which theory is belived to have higher impact.

1.2. Bacterial species and speciation driving forces

What are bacterial species? This question is far from simple and straightforward to be answered; it evokes a lot of discussions and disagreements among microbiologists, same as the issue of speciation and its driving forces. Below different approaches and concepts are described.

1.2.1. Phenetic species concept

Currently, a polyphasic (phenetic) species concept is being used by microbial taxonomists and systematists. Polyphasic taxonomy arose more than three decades ago, and it aims to integrate different kinds of data (phenotypic, genotypic, and phylogenetic) on microorganisms (Colwell, 1970; Vandamme *et al.*, 1996). Nowadays some concrete rules for the species affiliation exist: the species is estimated as a basic unit of taxonomy (Wayne *et al.*, 1987) and is “a group of strains, including the type strain, sharing 70% or greater DNA-DNA relatedness with 58°C or less D*Tm* (D*Tm* is the difference in the melting temperatures in degrees Celsius between the homologous and heterologous hybrids formed under standard conditions)” (Wayne *et al.*, 1987). Also, 16S rRNA sequence comparison might serve as the first indication that a new species has been isolated (less than 97% sequence similarity over full pairwise comparisons) (Tindall *et al.*, 2009).

According to the phenetic concept, phenotypic and chemotaxonomic characteristics of the bacteria should be also determined and must be in agreement with the results of the genomic comparison. A type strain must be denominated, and henceforth this strain is used as a reference specimen and name bearer (Staley & Krieg, 1984). In the current genome era propositions of new, several genome sequence based algorithms for the species affiliation emerge, which are meant to replace the laboratory work of DNA-DNA hybridization (Auch *et al.*, 2010; Richter & Rossello-Mora, 2009). In principle, values for the similarity of genome sequences are obtained by different algorithms and are calibrated to the 70% DNA-DNA hybridization threshold value for species delineation (Auch *et al.*, 2010; Konstantinidis & Tiedje, 2005; Richter & Rossello-Mora, 2009).

The polyphasic microbiological species concept partially matches to the eukaryotic species concepts. It is also phenetic, phylogenetic and genotypic. Phenetic in the way that it is numeric and attempts to classify organism basing on their overall similarity: on their morphological or other observable characteristics (though some of these traits are not necessarily universal for the whole taxon). It is phylogenetic in the sense that all the members of the group are closely related to each other and are descendants of the common ancestor. And, finally, polyphasic concept is based on the genotypic cluster definition (Mallet, 1995). There are definite numerical borders for genome comparisons,

which should guarantee allied interrelation of the members of the cluster (Rossello-Mora & Amann, 2001).

The polyphasic concept undoubtfully has a number of advantages; it is pragmatic and can be applicated to nearly all prokaryotes. The prokaryotic world is much more diverse than the eukaryotic world (Whitman *et al.*, 1998). Thus, the polyphasic approach is successfully solving the problem of the bacterial species affiliation, dealing with the enormous amount of diversity (Tindall *et al.*, 2009). In order to settle common rules for validly publishing the names of any new species, the Bacteriological Code of Nomenclature was arranged (Lapage *et al.*, 1992; Tindall *et al.*, 2009).

However, the current taxonomical approach is far from being perfect. It resulted in groups (species) that vary tremendously in their genetics, physiology and ecology (Cohan & Perry, 2007; Staley, 2006). Consequently, such situation led to numerous arguments and discussions between bacteriologists. Moreover, arguments start already at the point of what bacterial species should be understood as. Whether species should just represent phenotypically and genetically similar group of organisms? Or should it cluster together individuals with specific genomic/genetic, ecological, phylogenetic and evolutionary traits? In principle, microbiologists with an evolutionary but not a systematist's background, question the validity of the current taxonomic practice to capture microbial diversity from a truly biological perspective. Therefore, researchers have provided and continue to provide multitudinous attempts to propose the most appropriate and functioning way of species identification and affiliation.

1.2.2. The Ecotype concept

Frederick M. Cohan, Wesleyan University, CT, is one of the evolutionary biologists who efforts to set species definition and traits of its evolution. From his perspective, species determination in taxonomy cannot be separated from the evolution and ecology. Together with colleagues he developed a concept based on bacterial taxa, which are referred to as "ecotypes" (Cohan, 2001; Cohan, 2002; Cohan & Perry, 2007; Cohan & Koeppel, 2008; Koeppel *et al.*, 2008). In this concept authors combine ecological and evolutionary theory for the identification of the fundamental units of bacterial lineages (ecotype) via attempts to identify specific DNA sequence clusters. Ecotype is believed to be the group of bacteria that are ecologically similar to one another, so similar that genetic diversity within the ecotype is limited by a cohesive force, either periodic selection or genetic drift, or both (Cohan & Koeppel, 2008). It was demonstrated that commonly recognized nowadays named species are often in fact mixture of different ecotypes. Every each of which express their own properties of a species.

In order to get an experimental support for such a framework, ecotypes within the natural community of *Bacillus* clades sampled from the ‘‘Evolution Canyons’’ in Israel and Death Valley (USA) (Connor *et al.*, 2010; Koeppel *et al.*, 2008) were identified. For this, a sequence-based approach (‘‘Ecotype Simulation’’) was introduced to model the evolutionary dynamics of bacterial populations. As a result of the Ecotype Simulation, several ecotypes within the same named species were identified, each representing ecologically distinct lineage with specialization to different canyon slopes with the different solar exposure.

This approach is claimed to provide a lacking natural foundation for microbial systematics and ecology. To systematize and schematize the ecotype concept approach in agreement with the phenetic concept, it was suggested to add an “ecovar” epithet to the species binomial whenever several ecotypes are recognized within the admitted spectrum of a currently established species. Furthermore, in those cases where an ecotype is discovered to be outside the phylogenetic range of the species, such an ecotype should represent a new species (Koeppel *et al.*, 2008). The conclusion was made that such two folded approach should enrich bacterial systematics with ecologically significant, but previously ignored groups, while still respecting the stability of taxon names. Moreover, this approach would not violate any rules of the Bacteriological Code.

Some bacteriologists do welcome the incorporation of “evolutionary thinking” in the prokaryotic taxonomy at the level of species. Several articles were issued where the debate that “microbiologists need, in general, take a more natural view of the organisms they study” was raised (Ward, 1998; Ward *et al.*, 1998; Ward *et al.*, 2006), with the claim that it is time already to start paying attention to the evolutionary and ecological patterns of the species, and that bringing a more natural view to the prokaryotic species concept is of a the vital necessity. After study of the microbial biodiversity of the hot spring cyanobacterial mat communities by 16S rRNA sequences, the conclusion was drawn that in particular this method might be inappropriate for defining all ecologically specialized populations, as 16S rRNA sequences might be excessively conserved to succeed with such a goal (Ward *et al.*, 1998).

Another work dealing with the marine microbes’ communities attempt to identify and recover speciation driving forces of these bacteria (Polz *et al.*, 2006). Genome analysis revealed some variations among studied bacterial genome from given populations. And, according to authors, these variations were persistent with the lifestyles of every each population. Some cluster formations were observed, the same as distinct patterns of different loci and genome regions, allowing to conclude that they might be a result of intricate interaction of homologous recombination and selection (Polz *et al.*, 2006). In the

successive research that was provided with the use of the AdaptML [a maximum-likelihood-based framework tool for studying both the sequence evolution and ecological history of a set of gene sequences (Hunt *et al.*, 2008)] ecological differentiation of bacterioplankton and populations specific for seasons and life-styles (combinations of free-living, particle, or zooplankton associations) were discovered (Fraser *et al.*, 2009; Hunt *et al.*, 2008). Though these populations were mostly consistent with the named species, ecologically different populations on distinct levels of phylogenetic differentiation were uncovered for *Vibrio splendidus*. Authors believe that such a founding might be an evidence for recent and perhaps still ongoing adaptive radiation. This way it appears as environmental specialization (natural selection) might be of an important interrelation, or even serve as a trigger for speciation among sympatric bacteria (Fraser *et al.*, 2009; Hunt *et al.*, 2008).

1.2.3. Recombination as cohesive force in speciation

Nevertheless, there are a lot of argues and disagreements with the ecotype approach in which cohesive and driving speciation forces in bacterial species microevolution are pointed out to be either periodic selection or genetic drift, or both. In some works the main cohesive force is pointed out to be recombination, while genetic drift is believed to have rather a splitting power (Feil & Spratt, 2001; Fraser *et al.*, 2005; Fraser *et al.*, 2007; Hanage *et al.*, 2006a; Hanage *et al.*, 2006b; Spratt & Maiden, 1999). While analyzing population structure of some pathogenic bacteria (which is believed to represent fairly all genetic diversity), the population structure and molecular characteristics of pathogens (i.e. *Helicobacter pylori, Neisseria gonorrhoeae, Salmonella enterica and Neisseria meningitidis*) was determined and the conclusion was made that simple evolution model that is based on neutral mutational drift, modulated by recombination is the most appropriate for the microevolution description of these bacteria (Fraser *et al.*, 2005). Moreover, it was shown that the rate of the recombination has the strongest influence on the genetic discrepancy, and, thus, distinction should be made between clonal divergence and sexual speciation (Fraser *et al.*, 2007).

1.2.4. Geographical isolation driving evolutionary divergences (i.e. genetic drift)

Another driving parameter for the species microevolution was suggested to be physical isolation of populations by their geographical separation (Whitaker, 2006). For example, in a detailed study of the island-like hot spring cyanobacterial communities some divergence was discovered and explained by geographical isolation (i.e. genetic drift) and was pronounced to influence the evolution of bacteria (Papke *et al.*, 2003; Papke & Ward,

2004). Furthermore, after analysing the archaeon *Sulfolobus* by Multi Locus Sequence Typing (MLST) a similar conclusion was made: populations of hyperthermophilic organisms have recently diverged due to the geographical barriers, and even more, this relation was shown to be congruent on a global scale (Whitaker *et al.*, 2003). Even more, when MLST analysis of the tick-borne bacterium *Borrelia burgdorferi* (the causative agent of Lyme borreliosis) was performed, North American and European populations were revealed to be genetically different, this findings was also interpreted to be a result of geographical isolation (Margos *et al.*, 2008).

1.2.5. Do species exist at all?

Finally, one of the most discrete species concepts is the concept of “no species”. The main idea of this concept is that there is no such a thing as “species” in the microbial world (Bapteste *et al.*, 2009; Doolittle & Papke, 2006; Doolittle, 2008; Doolittle, 2009; Doolittle & Zhaxybayeva, 2009). Taxonomists, systematics and microbiologist according to this theory are dealing with the “man-made” names, in the nature there are no (and should not be) essential reasons for appearing any groups of individuals possessing coherent genetic and phenotypic features that would allow them to fit in the definition of “species” (Doolittle & Papke, 2006; Doolittle, 2008; Doolittle, 2009; Doolittle & Zhaxybayeva, 2009). Prokaryotic evolution and the tree of life are two different things, and they need to be treated such, and not extrapolating from macroscopic life to prokaryotes (Bapteste *et al.*, 2009). Lateral gene transfer and recombination are the processes that make it nearly impossible to come out with single species concept model. Any genome clusters that have been studied and described so far, might be explain by the pure chance or, rather by the limitations in culturing methods, samplings and/or used methods of genome diversity studies. Accepting the whole idea of natural species existence for granted, is the mistake that many of the researchers make and, therefore, they do accept any level of clustering to be another evidence of subsistence of “genuine” species (Doolittle, 2008; Doolittle, 2009; Doolittle & Zhaxybayeva, 2009). However, Doolittle and colleagues do not fully repel taxonomic practice (Dipippo *et al.*, 2009), as humans, perhaps as result of their evolved psychology, will always need to make classifications, as scientists, however, there is no need to believe in taxonomic classifications (Doolittle, 2009).

1.3. Population structure and style of microevolution of bacterial species

Microevolution of different bacteria has been under the high interest by the microbiological community for quite a while already and some meaningful information on this subjected was produced. Also, some attempts to analyze and classify the

accumulated insights were made. One of such efforts resulted in the detailed overview and graphical description of possible diversity of the evolutionary models (Cohan & Perry, 2007; Ward *et al.*, 2008). These models accounts for various ways in which diversity might be introduced into the population (i.e. via mutation, recombination, etc.), and how the inhabited niche might influence this diversity (i.e. via genetic drift, natural selection, geographic isolation, etc.). Among listed are following models: stable ecotype, genotype + boeing, genetic drift model, speedy speciation, species-less, nano-niche model, recurrent niche-invasion, cohesive recombination model and animal-like model.

Stable Ecotype model. According to this model, natural selection is the power that provides cohesion to the bacterial cluster. Within species diversity is proposed to be supplied by power of mutation and recombination. Ecotypes (ecologically distinct populations) hold different ecological niches, and are adapted to the occupied niche. Rather weak recombination do not allow big within-population diversity, and periodic selection events furthermore are able to purge the population by assuring that the most fitted in this environmental conditions mutant or recombinant sweeps the whole population. When one population invades a new ecological niche, separation process starts by natural selection force. As a result, two different phylogenetic lineages evolve, each representing ecologically different population (Cohan & Perry, 2007; Ward *et al.*, 2008).

Examples of microbial groups, possessing such a microevolution style, are *Mycobacterium tuberculosis* (Smith *et al.*, 2006b) and *Bacillus simplex* (Sikorski & Nevo, 2005).

Geotype-plus-Boeing model. In this model, cladogenesis is triggered by the physical (geographical) isolation and within population diversity is provided by accumulation of the neutral mutations.

Geographical isolation in the modern world is not as stable as it was previously. Microbes are transferred from place to place, from country to country and from the continent to continent together with people. Therefore, sometimes several bacterial distinguished groups can be found within one ecotype at one place, as can be observed with the bacterium *Yersinia pestis* (Achtman *et al.*, 2004). Thus, the hypothesis is that before the time of speedy human transportation, geographically separated and isolated populations of the same ecotype might have diverged into secluded clusters. With the advent of rapid and worldwide transportation systems, the endemic geotypes within single ecotypes start to merge geographically. As a result, separate populations, which nevertheless belong to the same ecotype, can now be observed in the same geographical location. So, the example of microbial group for the Genotype + Boeing model is *Yersinia pestis* (Achtman *et al.*, 2004).

Genetic drift model. This model mostly likely is represented in obligate pathogenic and endosymbiotic bacteria. The size of the population is rather small and limited, therefore, random genetic drift have very high impact and allow fixating allelic variants, resulting in subclusters of closely related individuals. Bacteria illustrating genetic drift model might be *Buchnera* (Herbeck *et al.*, 2003; Moran & Mira, 2001; Moran *et al.*, 2002).

Cohesive recombination. In the bacteria, in which genetic recombination rates are very high, recombination serves as cohesive force. Such a situation might be the case, when repetitive recombination events are observed within and between several closely related populations.

Helicobacter pylori can serve as a bacterial example of such a speciation model (Falush *et al.*, 2001).

Species-less model. In a rapidly changing environment new populations arise and extinct also very fast. Therefore, the genetic diversity within the species must not be limited into the specific habitat conditions, no evolution force must serve as a cohesive; vice versa, every each mutant or recombinant might serve as a founder for the new, rapidly inventing and extincting, species. Thus, the only cause of homogeneity within the cluster is the young age of closely related populations (species). Species-less model is proposed for being an appropriate in such situations.

Recurrent niche invasion. This model is an appropriate one in the species that iteratively gain or loose their specific niche adaptations (like easygoing gain or loss of a plasmid) and by this might gain adaptations for different ecological niche. In the lineage it results in repeated float from niche to niche, such a trend means that every population is not an irreversibly separate lineage, and might never become a separate cluster.

The best way to explain this model of speciation in details is by giving an example of a bacterium that performs it. *Bacillus thuringiensis* lineages can contain “alternative crystal toxin” plasmids, every of which is able to kill a different order of insect (Rasko *et al.*, 2007). This lineage while loosing one crystal toxin plasmid and then getting another, transfers from one ecological niche to another. If such situation appears to be frequent, then this lineage is not inconvertible separate sequence cluster, and might never become such. Another organism that supposedly examples this model of microevolution is *Rhizobium* (Mutch & Young, 2004).

Nano-niche model. In this model several lineages within single species exist. Every each of these lineages is adapted in a very slight way to their own nano-niche, within a single macrohabitat, they even might have their own periodic selection events [a ‘speciation-quashing’ event (Cohan & Perry, 2007)]. Still, one adapted lineage might outcompete and extinct other ecological lineages. These subgroups are not necessarily

inconvertibly separated; they rather do not have enough time to diverge into separate sequence clusters. This is exactly the behavior that possess *Vibrio splendidus* in its adaptation to the small particles in the marine water column (Thompson *et al.*, 2005).

1.4. Introduction into *Pseudomonas aeruginosa*

The first one to mention this bacterium in literature was Schroeter in 1872, in his publication it was called *Bacterium aeruginosum* (Schroeter, 1872). Since that time the name was changed several times, until the current name appeared and was validly published (Skerman *et al.*, 1980). Since the isolation, a lot of different information was gathered on its physiology, genetics, ecology and phylogeny.

It is a member of the *Gammaproteobacteria* class of Bacteria. Phylogenetically and taxonomically the species belongs to the family *Pseudomonadaceae*, which includes 12 species, with *Pseudomonas* as type species. The genus *Pseudomonas* consists of 188 species (Euzeby, 1997) and *P. aeruginosa* is the type species for this group (Olsen *et al.*, 1994).

Physiological characteristics of the *P. aeruginosa* are the following: it is a Gram-negative aerobic rod measuring 0.5 to 0.8 µm by 1.5 to 3.0 µm. Furthermore, it secretes different pigments like pyocyanin (blue-green), fluorescein (yellow-green and fluorescent), and pyorubin (red-brown) (Etymology: L. fem. adj. *aeruginosa*, full of copper rust, verdigris, hence green.). Preliminary it can be identified by its specific odor in vitro and by the blue-greenish color of the fluorescent producing colonies. The optimum growth temperature is 37°C; maximum temperature of growth is 42°C. Usually, the strains are motile by the means of a single polar flagellum.

The genome size of *P. aeruginosa* ranges from around 5.2 Mb (Hector & Johnson, 1990) till 6.3 Mb (Stover *et al.*, 2000a). It consists of a conserved core and a variable accessory part. The genetic elements are non-randomly distributed across the genome (Wiehlmann *et al.*, 2007). At the moment, seven genome sequenced strains (all are from clinical origin) are publicly available in the databases (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi; status May 2010).

Ecologically, *P. aeruginosa* is widespread in natural habitats, and can be isolated from the multiple environmental niches, such as water, soil, plants, animals and humans (Goldberg & Pier, 2000). *P. aeruginosa* is also an important clinical agent, as it is an opportunistic pathogen that causes wide range of acute and chronic injures and diseases of plants, animals and, very widely, humans (Chugani & Greenberg, 2007).

1.4.1. The distribution of *P. aeruginosa* in the environmental habitats

P. aeruginosa is a pervasive bacterium which invades a wide range of ecologically different habitats all over the world. It was widely reported to be isolated from different water sources: river water (Pirnay *et al.*, 2005), sea water (Kimata *et al.*, 2004) and opened ocean (Khan *et al.*, 2007; Khan *et al.*, 2008; Mena & Gerba, 2009). Also, *P. aeruginosa* is quite common in animals (Ozaki *et al.*, 1990) and animal feces (isolation by M. Vives-Flórez from the horse stables, personal communication) and soil and vegetable material (Green *et al.*, 1974). Besides these rather common habitats, *P. aeruginosa* is able to thrive also in extreme habitats such as sodium dodecyl sulfate (SDS) (Hagelueken *et al.*, 2006).

High ecological value at least for some strains of this bacterium were also reported (Bano & Musarrat, 2003; Hasanuzzaman *et al.*, 2004; Szoboszlay *et al.*, 2003). Some strains were proven to be good oil degraders (Hasanuzzaman *et al.*, 2004; Szoboszlay *et al.*, 2003). Other representatives of the *P. aeruginosa* were suggested to act as plant-growth promoting and biocontrolling rhizobacteria (Anjaiah *et al.*, 2003; Bano & Musarrat, 2003). Moreover, it produces glycolipidic surface-active molecules (rhamnolipids) which have potential wide-range biotechnological applications (Pham *et al.*, 2004; Soberón-Chávez *et al.*, 2005).

1.4.2. The clinical significance of *P. aeruginosa*

P. aeruginosa is able to cause a large set of different infections, both acute and chronic; it can invade any part of the body (http://emedicine.medscape.com/article/226748-overview; status 22 June 2010). It commonly infects the respiratory tract of immunocompromised patients and people with the cystic fibrosis disease (CF). In individuals, abusing intravenous drugs, *P. aeruginosa* might infect native heart valves (Bicanic & Eykyn, 2002). Invading central nervous system *P. aeruginosa* causes meningitis and brain abscess (Davidson *et al.*, 1982). Patients with uncontrolled diabetes are in a group risk for getting chronic otitis media; furthermore, extension of this otitis can result in osteomyelitis, and, more dangerously, can create cranial nerve palsies (Sander, 2001). Eye-associated infections result in bacterial keratitis and scleral abscess (Hobden *et al.*, 1999). *P. aeruginosa* is also able to involve in vertebral column, pelvis, and sternoclavicular joint, causing vertebral osteomyelitis (Watanakunakorn, 1975). *Pseudomonal* urinary tract infections are mostly hospital-acquired and typically appear after catheterization, instrumentation, or surgery (Shigemura *et al.*, 2006). If patients suffering from neutropenia acquire *P. aerugionosa*, typhlitis is most likely to be developed (http://emedicine.medscape.com/article/226748-overview; status 22 June 2010). And,

finally, *P. aeruginosa* is one of the “curses” of the burn units where it is one of the most frequent source of burn wound sepsis (Baltch, 1994).

One of the biggest problems in curing infections associated with *P. aerugionosa* is that this bacterium often prospers in clinical environments. There are multiple reports of detecting multi-resistant strains of *P. aeruginosa* from hospital floors, bed rails, sinks and from the hands of medical personnel [i.e. (Altoparlak *et al.*, 2005; Chitkara & Feierabend, 1981; Hsueh *et al.*, 1998)]. Due to the transfer from patient to patient, multi-drug resistant clones can remain in hospitals for many years (Hsueh *et al.*, 1998), and, thus, it is one of the major nosocomial pathogens which attribute to the high percentage of patient mortality and morbidity (Altoparlak *et al.*, 2005).

1.5. Population structure of *P. aeruginosa*

Most population genetic studies have been performed mainly under clinical perspectives (Pirnay *et al.*, 2002; Pirnay *et al.*, 2009; Rakhimova *et al.*, 2009; Wiehlmann *et al.*, 2007). Analysis of the genome linkage revealed that the recombination rate varies considerably in relation to the chromosomal location (Romling *et al.*, 1995) and genomic regions, encoding housekeeping, provide low rates of recombination, whereas loci, encoding accessory genes, have higher rates of recombination and, thus, might be subjected to the comparable higher selection rates then housekeeping genes (Romling *et al.*, 1995). Furthermore, the apparent consensus in the literature is that *P. aeruginosa* has a non-clonal epidemic population structure (Curran *et al.*, 2004; Denamur *et al.*, 1993; Picard *et al.*, 1994; Pirnay *et al.*, 2002; Pirnay *et al.*, 2009). Epidemic structure assumes the existence of highly successful epidemic clones or clonal complexes, and indeed, several of these epidemic clones could be identified, such as the famous clone C or the clinical MDR serotype O:12 lineage (Giammanco *et al.*, 1985; Pirnay *et al.*, 2009; Watine, 1999).

It is frequently stated that there is no difference between *P. aeruginosa* strains from different environmental or clinical habitats (Alonso *et al.*, 1999; Foght *et al.*, 1996; Kiewitz & Tummler, 2000; Morales *et al.*, 2004; Ruimy *et al.*, 2001). However, such notion appears to be questionable already in the light of the widely accepted observation that *P. aeruginosa* strains, once they have been occupied into a clinical habitat such as the human body, readily evolve in the due course by genetic changes in order to microevolutionary adapt to the characteristics of the human body (Lee *et al.*, 2006; Smith *et al.*, 2006a). Mutated factors include the loss of O-antigen components of the lipopolysaccharide (LPS), loss of twitching motility, loss of pyoverdine production, loss of secreted protease activities (including elastase), and reduced biofilm formation (Smith *et*

al., 2006a). Changes of other genes required for pathogenicity are not possible to follow, as virulence in *P. aeruginosa* is proposed to be both multifactorial and combinatorial, genes that might be responsible for pathogenicity in one strain, are “neither required for, nor predictive of virulence in other strains” (Lee *et al.*, 2006).

1.6. Phenotypic characterization of *P. aeruginosa*

Little is known about the wide spectrum phenotypical characterization of environmental isolates of *P. aeruginosa.* Clinical isolates have been always under close attention, thus, there are some data on the physiological and phenotypical profile changes of the clinical strains of *P. aeruginosa*. Most common changes are the ones that are associated with its virulence in acute infections. Alterations include conversion to mucoidy, loss of flagella or pilus effected motility types, loss of serotype-specific antigenicity, and advent of the antibiotic multiresistant strains (Barth & Pitt, 1995; Luzar & Montie, 1985; Luzar *et al.*, 1985; Makin & Beveridge, 1996; Pedersen *et al.*, 1986a; Pedersen *et al.*, 1986b; Sener *et al.*, 2001). Possible explanation of physiological changes might be associated with the adaptation processes allowing bacteria to survive in the hostile surrounding of the lungs of CF patients (Drenkard & Ausubel, 2002; Mahenthiralingam *et al.*, 1994; Mahenthiralingam & Speert, 1995).

Motility. *P. aeruginosa* utilizes three ways of cellular motility: swimming, swarming and twitching; the choice for the movement type depends on the situation (Fraser & Hughes, 1999). Swimming and swarming are guaranteed by flagella, twitching is being mediated by type IV pili. Planktonic *P. aeruginosa* residing in fluidic environments (river, sea, ocean, etc.), displays a swimming motility type. When cells are on semi-solid surfaces, or solid surfaces with the liquid layer, swarming is the type of performed movement, cells transfer to the hyperflagelated and elongated phenotype and start to move in a coordinated maner. Being on solid surfaces, *P. aeruginosa* cells, by means of pilus retraction and extension, move in a way called *twitching motility*. Twitching is necessary for infectivity and biofilm formation process (Beatson *et al.*, 2002a; Beatson *et al.*, 2002b; Fraser & Hughes, 1999; Leone I., 2008; Rashid & Kornberg, 2000).

Biofilm formation. Biofilm is the form of bacterial communities which are adhered to a surface and it is believed that mostly bacteria are indwelled in the environment in this form (Costerton *et al.*, 1995). It protects bacteria from hostile influences from the habitat, such as UV exposure, metal toxicity, acid exposure, dehydration and salinity, as well as infection challenges (i.e. phagocytosis, antibiotics and antimicrobial agents) (Stoodley *et al.*, 2002).

In response to different signals, such as nutritional and/or oxygenic status of habitat, individual planktonic cells initiate interaction with surface and form intricate and highly structured bacterial communities which are termed biofilms (O'Toole & Kolter, 1998a; O'Toole & Kolter, 1998b; Pratt & Kolter, 1998; Wimpenny & Colasanti, 1997). Fully developed biofilms can perform different architectural and physical/chemical qualities (Costerton *et al.*, 1995). Some cells of *P. aeruginosa* within such a biofilm were shown to behave differently from those that exist in planktonic form, providing different genes expression profile for some genes [i.e. *algC* gene required for the synthesis of extracellular polysaccharides (Davies *et al.*, 1993; Davies & Geesey, 1995) and *ampC* gene which encodes for a beta-lactamase precursor (Bagge *et al.*, 2004a; Bagge *et al.*, 2004b)] and phenotypes, at least with the respect to the heightened antibiotic resistance (Hoyle BD, 1991). Although it is believed that most (though not all) of the cells present in the biofilm reside in the stationary physical phase (Waite *et al.*, 2005; Waite *et al.*, 2006), it was shown that some adaptations are necessary for the, i.e., anaerobic oxidation and iron-limitation stress (Hentzer M, 2003).

Several factors are known to be needed and involved in the biofilm maintenance and development of *P. aeruginosa*, such as flagella, type IV pili and *las-rhl* quorum-sensing system (Davies *et al.*, 1998; O'Toole & Kolter, 1998a; O'Toole & Kolter, 1998b). This model for *P. aeruginosa* biofilm formation, presenting involvement of flagella in attachment, the need for twitching ability (which is in *P. aeruginosa* type IV pili-driven motility) in microcolony formation and the following development of microcolonial structures, is generally accepted to be common for *P. aeruginosa* biofilm development (Costerton *et al.*, 1999; O'Toole *et al.*, 2000; Stoodley *et al.*, 2002).

Pyocyanin production. *P. aeruginosa* strains are able to produce two types of soluble pigments: pyoverdine, the fluorescent pigment and the blue pigment pyocyanin. Pyocyanin is produced by cells after the exponential phase of growth. This blue pigment is highly important for the *P. aeruginosa* due to the ability to provide antibiotic activity against a wide variety of bacteria (Baron & Rowe, 1981; Hassan & Fridovich, 1980; Schoental, 1941), to the capacity to kill mammalian cells (Hassett *et al.*, 1992) and to raise apoptosis level of human neutrophils (Hassett *et al.*, 1992). Considering its biological activity, the intensive production of pyocyanin by clinical isolates of *P. aeruginosa* was detected consistently (Finnan *et al.*, 2004; Fothergill *et al.*, 2007; Reyes *et al.*, 1981; Wilson *et al.*, 1988).

Pyocyanin's cytotoxicity is strongly linked to its ability to redox (reduction-oxidation reaction) cycle (Hassan & Fridovich, 1980). It accepts electrons directly from either NADH or NADPH and then, under aerobic conditions, it passes them further to O_2, which leads to

the generation of the reactive oxygen species (ROS) superoxide (O_2^{--}) and hydrogen peroxide (H_2O_2) (Hassan & Fridovich, 1980). Furthermore, with the use of cellular reducing equivalents (e.g., NADH or NADPH), addition of pyocyanin to cells places them under increased oxidative stress (Hassan & Fridovich, 1980).

Proteolytic activity. *P. aeruginosa* can secrete alkaline protease and two elastases, which have been distinguished as exoenzymes and virulence factors (Howe & Iglewski, 1984; Kernacki *et al.*, 1995; Kessler *et al.*, 1993; Ohman *et al.*, 1980; Twining *et al.*, 1993). Proteases produced by *P. aeruginosa* were shown to be important for the establishment of an infection, as well as for the amount of damages caused by it (Baker, 1982; Nicas & Iglewski, 1985). Moreover, protease activity was demonstrated to play a role in tissue damaging and enhancement in virulence in different human infections, such as damaging respiratory epithelium of cystic fibrosis patients (Suter, 1994) and increased virulence of *P. aeruginosa* in burn patients (Pavlovskis & Wretlind, 1979).

Survival in water. There are numerous reports existing on the survival of *P. aeruginosa* in water (Byrd *et al.*, 1991; Favero *et al.*, 1971; Kayser W. V. *et al.*, 1975; Khan *et al.*, 2007; Legnani *et al.*, 1999; Sjogren & Gibson, 1981). Different types of water were tested in these studies: drinking water, clinical distilled water, natural mineral water, lake water, etc. In spite the fact that reported survival periods varies quite substantially [from 95 days (Byrd *et al.*, 1991) up to 5 years (Legnani *et al.*, 1999)], the same conclusion was continuously made: *P. aeruginosa* cells are able to survive durably in different aquatic surroundings (Byrd *et al.*, 1991; Favero *et al.*, 1971; Kayser W. V. *et al.*, 1975; Legnani *et al.*, 1999; Sjogren & Gibson, 1981).

Serotypes. *P. aeruginosa* is known to be serologically heterogeneous, the heat-stable lipopolysaccharide (LPS) component of the outer membrane determines the O-antigenic type and the core LPS components are conserved throughout the species (Chester *et al.*, 1973). Numerous serological typing schemes have been described for *P. aeruginosa* (Dayton *et al.*, 1974; Fisher *et al.*, 1969; Habs, 1957; Homma *et al.*, 1971; Lanyi, 1966; Meitert, 1964; Sandvik, 1960; Verder & Evans, 1961) but only a few have been widely used (Fisher *et al.*, 1969; Habs, 1957; Homma *et al.*, 1971; Verder & Evans, 1961). All these schemas were not 100% comparable to each other and the necessity for the unitary system arose. Following such urge, the Subcommittee on *Pseudomonadaceae* and Related Organisms of the International Committee of Systematic Bacteriology of the International Association of Microbiological Societies (IATS) proposed seventeen O-groups as the standards for serotyping *P. aeruginosa* (Liu *et al.*, 1983).

Though the serotype profile was thought to be a rather stable characteristic of *P. aeruginosa* strains, polyagglutination, autoagglutination and “no agglutination” phenomena

were shown to be quite frequent for clinical isolates, especially for strains that induced chronic infections (Hancock *et al.*, 1983; Pitt *et al.*, 1986). Polyagglutinating strains, are those that get agglutinated by more than one antiserum towards unrelated O-antigens; autoagglutinating strains are those that get self-agglutinated in the physiological saline in the absence of antiserum; and, finally, “no agglutination” phenomena is observed in strains that do not get agglutinated by any antiserum (Hancock *et al.*, 1983). Independent reports have confirmed that these serologically atypical forms are antigen variants which are defective in their ability to synthesize O-specific side chains of LPS (Hancock *et al.*, 1983; Pitt *et al.*, 1986).

Bacteriophages. Bacteriophages are the most frequent and ubiquitous biological entities in our world – they are everywhere: in the oceans, soil, deep sea vents, the water we drink, and the food we eat (Sulakvelidze & Kutter, 2005). Their total count is estimated to be in the range from 10^{30} to 10^{32} in total, and their essential role in every ecosystem (in which this ever was explored) was proved to lie in regulating the microbial balance (Sulakvelidze & Kutter, 2005).

Lately, the increase in interest for bacteriophages’ studies can be observed. Such a expansion of this field of research is happening due to several reasons - microbiologists started to recognize phages’ ecological essence, their vital role in life’s origin and evolution, and their therapeutic potential (Ceyssens *et al.*, 2009). The therapeutic potential of bacteriophages is becoming more and more promising and a vital field of study, since nowadays we are experiencing an era where the problem of antibiotic resistance becomes so substantial and widespread (O'Flaherty *et al.*, 2009). *P. aeruginosa* is one of the organisms that become problematic to be successfully treated - more and more often the antibiotic treatments fail because an increasingly large number of strains becomes multiresistant towards antibiotics (Palmer *et al.*, 2005; Sadikot *et al.*, 2005; Sriramulu *et al.*, 2005).

1.7. Tools to analyze the genetic diversity among *P. aeruginosa* strains

The genome of the *P. aeruginosa* has a mosaic structure which consists of a highly conserved core, interrupted by variable accessory segments (Wolfgang *et al.*, 2003). The core genome is represented by synteny of multiallelic conserved genes and the accessory genome, which include a set of genomic islets and genomic islands (Wolfgang *et al.*, 2003).

Several commonly used methods of identification of the genetic diversity of bacteria, including *P. aeruginosa*, exist. Genome fingerprinting, such as Amplified Fragment Length Polymorphism / Restriction Fragment Length Polymorphism (AFLP/RFLP) and multiplex

PCR analysis (the simultaneous PCR-based generation of multiple differently sized DNA amplicons) often give quite big variations and, thus, are hardly comparable among laboratories. On another hand, MLST and MLEE methods which use nucleotide sequence data of, mostly, seven housekeeping genes, are rather universal, portable and definitive methods for bacteria characterization (Maiden *et al.*, 1998; Maiden, 2006). However, MLST method provides knowledge only on the genetic diversity of the core genome, moreover, currently available MLST database is rather poorly annotated with the respect of information on type of habitat and geography of isolation of strains and, thus, is of no use for the study of the influence of environment on the *P. aeruginosa* microevolution.

One of the latest offered methods for the genetic diversity affiliation is the method based on binary microarray (Wiehlmann *et al.*, 2007), which consists of 58 targets representing the core and the accessory genome. The core genome is analyzed by 13 single-nucleotide polymorphisms (SNPs) at seven conserved loci and two multiallelic loci (flagelin fliC and pyoverdine receptor fpvA). The accessory genome is tested with set of genetic markers that identify 10 genomic islets and six types of genomic islands (Wiehlmann *et al.*, 2007). Moreover, current database is fairly well big and annotated: it contains more than 1700 isolates from 658 independent sources (both clinical and environmental) with the information on the origin and isolation time on majority of strains.

Thus, the ArrayTube hybridization method (Wiehlmann *et al.*, 2007) is at the moment the most promising and appropriate for the presented study.

2. Project rationale

Despite the strong clinical focus, a result of the substantial pathogenic threat for humans, *P. aeruginosa* is widely and probably predominantly distributed as a free-living organism in water and soil (Gooderham & Hancock, 2009; Pirnay *et al.*, 2009). For example, clone C is widespread in natural environments (Romling *et al.*, 1994; Romling *et al.*, 2005). Thus, besides nosocomial transfer, the natural environment is obviously a common source of infection (Curran *et al.*, 2005; Scott & Pitt, 2004). However, only few studies explicitly focus on environmental strains (Khan *et al.*, 2008; Pirnay *et al.*, 2005), whereas in most population genetic studies the environmental strains are underrepresented (10-25%) (Pirnay *et al.*, 2002; Pirnay *et al.*, 2009). Nevertheless, though there is an obvious underrepresentation of environmental strains, most studies lean out to suggest clinical and environmental strains to be indistinguishable in terms of affiliation to (phylo-)genetic clusters or in phenotypes (Alonso *et al.*, 1999; Kiewitz & Tummler, 2000; Morales *et al.*, 2004; Ruimy *et al.*, 2001). Also, until now, the population genetic studies aimed mainly at identifying or exploring the nature of the *P. aeruginosa* population as such, which resulted in the notion of *P. aeruginosa* to have a non-clonal epidemic structure (Curran *et al.*, 2004; Pirnay *et al.*, 2002; Pirnay *et al.*, 2009). Yet, far less emphasis has been laid in further analysing the intrapopulation structure with respect to exploring the forces that may shape differently the observed different subgroups within *P. aeruginosa* and potentially foster their microevolutionary divergence.

The gap in knowledge on the population structure of environmental *P. aeruginosa* and on the reciprocal interaction of *P. aeruginosa* with the environment can be fulfilled only by providing a wide spectrum study, focusing not only on the genetic traits, but also on the phenotypic characterization of the natural population of *P. aeruginosa*, which should be samples for by a substantially amount of isolates. Thus, this PhD thesis was designed in an attempt to fulfill this gap in information.

3. Materials and Methods

3.1. Sampling places, sampling, and isolation of *P. aeruginosa*

The sampling regime was designed such to sample two river systems in Northern Germany (Oker/Aller and Weser) over a length of each 150-200 km at approximately uniformly dispersed sampling sites. Among a variety of potential sampling places only those were utilized which were under regular monthly water quality observation by the Lower Saxony Ministry of the Environment (Niedersächsisches Umweltministerium, www.mu.niedersachsen.de). The water load concentrations for ten chemical substance groups (total nitrogen, ammonium, nitrate, nitrite, adsorbable organic halogen compounds, phosphate, sulfate, total organic carbon, total phosphate and chloride; values are available from the Ministry on request) are summed by the Ministry into seven classes (I, I-II, II, II-III, III, III-IV, and IV), ranging from I (no load) to IV (very strong load). Class II is regarded by the Ministry to be the target-setting for an optimal water quality. For this study, we averaged the water load measurements from the approximately last five years in order to yield via an educated guess across all ten chemical substance groups three water quality categories (a) low pollution (classes I, I-II, II), (b) moderate pollution (classes II-III, III,), and (c) strong pollution (classes III-IV, IV) for the chosen sampling places. For detailed information on chemical concentrations per class, see Appendix for Supplemental Table 1.

The sampling sites were chosen such to cover approximately equal amounts of low, moderate, and strongly polluted sites. Altogether 13 sampling places were chosen (for more information on the geographical distribution of the sampling sites, see Supplemental Figure 10). The samplings were performed in October 2007 and July 2008. Each sampling was performed within two days. The samples were taken in direction of the water flow from the spring to the estuary. From each sampling station one liter of water was sampled in a sterile glass bottle and processed for strain isolation on the same day.

For this, the water was vacuum filtered through 0.45 µm mixed cellulose ester filter (90 mm in diameter) (Schleicher&Schuell; No. 5058945, Omnilab, Germany) (Pirnay *et al.*, 2005). Each liter of water was separated prior to filtration into 200 ml and 800 ml subsamples. After filtration, the filters were cut aseptically into three equally sized thirds, each of which was placed separately face up on agar plates with selective *Pseudomonas* medium (OXOID) as previously described (Pirnay *et al.*, 2005).This procedure should ensure that potentially identical isolates were not a result from laboratory enrichment but were present as such already in the environment. The plates were incubated for 36h at 37°C. Putative *P. aeruginosa* colonies, which provided a greenish-bluish color, were

further purified on selective *Pseudomonas* medium and subjected to a species-specific PCR targeting the *oprL* gene (De Vos *et al.*, 1997). The strains were artifiacially concentrated and stored in 15% glycerol at -80°C.

3.2. Molecular characterization of the isolates

ArrayTube hybridizations. The *P. aeruginosa* isolates were genetically characterized using the *P. aeruginosa* ArrayTube system as previously described (Wiehlmann *et al.*, 2007). Strains for biomass production were grown for 24h at 37°C on the *Pseudomonas* agar media (Oxoid). The method is based on binary microarray (Wiehlmann *et al.*, 2007), which consists of 58 targets representing the core and the accessory genome. The core genome is analysed by 13 single-nucleotide polymorphisms (SNPs) at seven conserved loci and two multiallelic loci (flagelin *fliC* and pyoverdine receptor *fpvA*). All targets are amplified from the *P. aeruginosa* colonies by the multiplex primer extension reaction with random labeling performed by incorporation of biotin-16-dUTP. The multiplex amplificate was hybridized under the high stringency with the oligonucleotide microarray of the target sequences that is inserted into the tip of a standard Eppendorf-like microtube. However, both the annealing temperature in the multiplex PCR and the hybridization incubation temperature were set from 60°C to 56°C in order to obtain clearer signals on the chip (Wiehlmann, personal communication). The hybridization signals are automatically converted to the multilocus genotype. The 16 binary SNP genotypes are represented by a four-digit hexadecimal code. The 16 SNPs are divided into four groups of four SNPs each, and the 16 possible combinations in each group are differentiated by 16 characters: 0–9, A–F (Wiehlmann *et al.*, 2007).

DNA extraction. Around 1.5 % percents of strains provided too mucoid biomass and could not be subjected directly for the ArrayTube hybridization. Thus, DNA extraction was performed prior hybridization. DNA was extracted by the means of the “Fast DNA SPIN for Soil Kit” (MP Biomedicals, LLC). DNA was extracted according to the manufacturer protocol. The final concentration of pure DNA in the PCR ArrayTube mix was 1.5-2 µg.

exoS/U multiplex PCR. In addition to the ArrayTube results the presence or absence of *exoS* and *exoU* genes was verified by a separate PCR as described previously (Ajayi *et al.*, 2003), with two minor modifications.

First, primers for only genes *exoU* and *exoS* were used:

exoS (118-bp fragment)

5'-Primer (position 686) .....ExoS-MP5, 5'-GCGAGGTCAGCAGAGTATCG-3'

3'-Primer (position 804) .....ExoS-MP3, 5'-TTCGGCGTCACTGTGGATGC-3'

exoU (134-bp fragment)

5'-Primer (position 1265) .....ExoU-MP5, 5'-CCGTTGTGGTGCCGTTGAAG-3'

3'-Primer (position 1399) .....ExoU-MP3, 5'-CCAGATGTTCACCGACTCGC-3'

Second, the reaction was run in a 2.5% Agarose High Resolution (ROTH, CARL ROTH GMBH + CO. KG, Karlsruhe) gel with 0.5 mg of ethidium bromide/ml.

In order to approach the whole genetic diversity of the *P. aeruginosa*, every each clone (haplotype) was subjected for the multiplex PCR; from 30% to 100% of the strains per clone were utilized for this analyses, depending on the size (amount of strains) and geographical diversity of the clone (in total a sum of 211 strains).

Agarose gel electrophoresis. Agarose gel electrophoresis was performed according to Sambrook et al. (Sambrook *et al.*, 1989). Each DNA sample was mixed with the 1/5 volume of loading buffer and loaded on a 2.5% (w/v) gel. The TAE buffer was used as a running buffer.

3.3. Media

Two types of media were used: "*Pseudomonas* selective Agar medium", OXOID (prepared according to the manufacturer protocol; Pseudomonas CN Selective Supplement that was used has an ordering number SR0102) and solid or liquid Luria Bertani Medium (LB).

One liter of LB: 10 g of Tryptone; 5g of Yeast extract; 10 g of NaCl, and 15 g of Agar (only for solid media).

3.4. Phenotypic characterization of the isolates

Biofilm growth, microscopy and data analysis. In order to quantify biofilm production, 5µl from a thawed glycerol culture were transferred into 150µl of LB within a 96-well plate. The pre-culture plate was covered with an air-permeable BREATHseal™ cover foil (Greiner Bio-One) and incubated at 37°C on a shaking unit. After 4h, 5µl of the pre-cultures of each well/strain were transferred to 100µl LB medium within a sterile half area 96-well µClear® microplate (Greiner Bio-One). The plates were sealed with a new air-

permeable cover foil and placed in an incubator with a humid atmosphere for biofilm development. The image-based screen was carried out after 70h of incubation. The incubation was paused after 24h in order to stain the bacteria with the LIVE/DEAD® *Bac*Light™ Bacterial Viability Kit (Molecular Probes/Invitrogen). 50µl of a diluted staining solution was added to each well resulting in a final concentration of 1.4µM Syto9 and 8.3µM propidium iodide (PI). Microscopy was performed with dual-headed, confocal microscope (Olympus Fluoview 1000) equipped with a 20x/0.75NA air objective (UPLSAPO) in the middle of each well. The z-step size was 2µm. Syto9 and PI were both excited with a 488nm Argon-Laser and emission detected around 515nm (Syto9) and 560nm (PI). The image-stacks obtained by fluorescence microscopy were converted to single bmp-images with the software Matlab® (version 7.5.0.342, The Mathworks) and further processed to obtain thresholded binary images using the Otsu thresholding-algorithm. Afterwards, stacks for each position and dye were reassembled with the software Auto PHLIP-ML (v1.0.0) (Merod *et al.*, 2007), now containing thresholded binary images instead of raw images. Analysis of the different stacks was performed with the Matlab implemented tool PHLIP (Mueller *et al.*, 2006) to yield the value for “biovolume”, which is the descriptive parameter for the physical properties of the biofilm.

Motility. In order to determine the swarming activity 1µl of a fresh over-night culture in LB broth was transferred using a custom 96-well pin device onto the surface of brain heart infusion agar plates containing 0.5% (wt/vol) Difco agar. Colonies were scored for differences in swarming motility compared to the “normal type” (the swarming size diameter produced by the majority of the environmental isolates) after 48h of incubation at room temperature. The swarming activity was semi-quantitatively classified into “minimal”, “normal” and “hyper” swarming strains. Similarly, for determination of swimming and twitching motility activity, 1µl aliquots of mid-log-phase cultures in LB broth were spotted onto LB agar plates with 0.3% (swimming) and 1% (twitching motility) agar (wt/vol) and incubated at 37°C. The diameters of the swimming zones (after 20h of incubation) and the twitching motility zones (after 24h and 48h) was semi-quantitatively classified into “non-motile”, “minimal”, “normal” and “hyper” swimming and into ”minimal”, “normal” and “hyper” twitching activity.

Proteolytic activity. In order to determine general proteolytic activity, 1µl of the fresh cell material from an LB culture was spotted onto 4% skim milk (Loewe Biochemica) agar plates and incubated for 24h at 37°C. Zones of clearance due to proteolytic activity were classified as ”no activity”, “minimal”, “normal” (the majority of strains) and “increased” activity.

Pyocyanin production. In order to determine the general level of pyocyanin production, 5µl samples of a thawed glycerol culture were transferred into 150µl of LB within a 96-well plate. The pre-culture plate was covered with an air-permeable BREATHseal™ cover foil (Greiner Bio-One) and incubated at 37°C on a shaking unit. After 24h levels of the different intensities of greenish-bluish color production of the pyocyanin was classified as “normal”, “increased” or “strong”.

Survival in water. LB broth was inoculated with 5µl from a thawed glycerol culture and incubated for 24h at 37°C. The culture was split, of which one half was washed 3x times with autoclaved distilled water (dH_2O) and the other half with autoclaved river water (from the sample station ”Peine”, which is regarded as strongly polluted) The washed cells were resolved in autoclaved dH_2O or the autoclaved river water, respectively, and adjusted to a cell concentration of 10^7 cells/ml. A volume of 1.5ml of each assay was then transferred into a single 4 ml flat bottom well on a 24-well culture plates (Corning Incorporated, USA) covered with an air-permeable BREATHseal™ cover foil (Greiner Bio-One) and incubated at 70 rpm at approx. 22°C (room temperature). Culturable cell counts were determined every 3-4 days by serial dilution plating on the *Pseudomonas* selective media. (OXOID). For a given assay, the experiment was terminated when no culturable cells were observed anymore in the remaining incubation fluid. In order to test for perhaps viable but not culturable cells remaining in the visually emptied well, the wells were further incubated with 1 ml LB broth. In none of the subsequent LB assays cell proliferation was observed, suggesting indeed the death of cells during the course of experiment. Two randomly chosen strains from both ext_cc groups were subjected to the duplicate experiment on the survival in distilled water.

Serotyping. Serotyping for O-group specific antigens was performed by using the “Antiserum Pseudomonas *aeruginosa* set” (BioRad) following the manufacturers protocol. First, samples were tested with the polyvalent sera: PMA, PME, PMC and PMF. Furthermore, monovalent sera were used - 16 monovalent sera numbered from 1 to 16. Polyvalent sera and monovalent sera were supplied in 3 ml dropper bottles (60 tests). Serotyping was performed on fresh pure cultures of *P. aeruginosa* grown for 24h at 37°C.

3.5. Phage assays

The phages were isolated and purified from the sewage plant Steinhof in Braunschweig. The phages JG003, JG004 and JG005 were isolated using strain PAO1, the phages JG024, JG025, JG026, and JG028 were isolated using strain PA14 (Julia Garbe, PhD thesis 2010, TU Braunschweig, Germany; accessible at http://rzbl04.biblio.etc.tu-bs.de:8080/docportal/servlets/MCRFileNodeServlet/DocPortal_derivate_00008650/Doktorarbeit.pdf;jsessionid=DF821E192BF37415A09826A43ADE2AB8). To determine the phage host range, top-agar LB plates were prepared by adding approximately 3.5×10^8 cells of *P. aeruginosa* from an overnight LB broth to 3.5 ml of LB top agar (0.75 %). Ten µl of a phage stock solution (approximately 10^8 phage particles per ml) were spotted on the top-agar plate and incubated at 30°C for 24h. The lysis zone at the site where the phage suspension was applied was classified as “complete lysis”, “semi lysis” and “no lysis” (Knezevic *et al.*, 2009). Each phage was tested against each bacterial strain in at least duplicate in independent experiments.

3.6. Statistical analysis.

The non-parametric Mann-Whitney U test (MWU) was applied for comparing the means of two groups. For comparison of the means of three or more groups, the novel and robust method “multcomp” was applied (Herberich *et al.*, 2010). All other statistical analyses were performed as indicated and referenced in the text.

4. Results

4.1. The characteristics of distribution of *P. aeruginosa* in Northern Germany river systems

P. aeruginosa was successfully isolated from all 13 sampling stations in winter and in 12 out of 13 stations in summer (Tab. 1), with two to 56 strains per station (detailed results on the strains' abundancy for every each sampling station and both seasons is available in the Appendix as Figure 12). A significantly larger number of isolates per sampling station was obtained in summer (p = 0.03; Mann-Whitney U test, MWU; column "as" in Tab. 1), also, the number of strains per clone is significantly higher in summer (column "s/cl"; p = 0.011, MWU).

Table 1: Characteristic values for the genetic structure of *P. aeruginosa* in different sampling stations and seasons

Station[1]	R[2]	E[3]	S[4]	Abundancies[5]				Haplotype configuration[6]	Neutrality test	Genetic diversity	Seasonal differences	
				as	s	cl	s/cl		(EW/S)[7]	π[8] (s/cl)	Abundance[9] (as/hr)	π[10] (s/cl)
Goslar	O	L	W	4	4	1	4.0	4^{sol}	n.a.	n.a.	S / W	n.a.
bei Börßum	O	M	W	8	8	5	1.6	3^{B} 2 1 1 1	ns / ns	6.21 / 7.6	S / W	S / W
Groß Schwülper	O	M	W	44	42	15	2.8	17^{sol} 8 4 2 (11×1)	* / **	6.45 / 7.66	W / S	W / W
Peine	O	H	W	28	18	15	1.2	8^{B} 4 4 (12×1)	ns / *	6.29 / 6.86	S / S	W / S
Celle	O	L	W	9	6	6	1.0	(6×1)	n.a.	7.33 / 7.33	S / W	W / W
Han.Münden	W	L	W	28	27	14	1.9	7^{S42} 4 3 2 2 (9×1)	ns / ns	5.33 / 6.18	W / S	S / S
Hes.Oldendorf	W	M	W	22	22	12	1.8	5^{B} 3 (4×2) (6×1)	ns / ns	6.96 / 7.40	S / W	W / W
Ströhen	W	H	W	15	15	11	1.4	(4×2) (7×1)	ns / ns	6.65 / 7.05	0 / W	W / W
Allerbruch	W	H	W	2	2	2	1.0	1 1	n.a.	8.00 / 8.00	n.a.	n.a.
Drakenburg	W	H	W	9	9	7	1.3	3 (6×1)	** / **	6.11 / 6.76	W / W	W / W
Langwedel	UW	L	W	12	12	11	1.1	2 (10×1)	n.a.	5.85 / 6.04	S / W	W / S
Brake-Weser	UW	H	W	7	7	6	1.2	2 (5×1)	n.a.	6.57 / 7.00	S / W	W / S
Brake-Sieltief	UW	H	W	2	2	2	1.0	1 1	n.a.	5.00 / 5.00	S / W	S / S
Goslar	O	L	S	11	10	2	5.0	9^{sol} 1	** / **	1.00 / 5.00		
bei Börßum	O	M	S	24	20	10	2.0	5^{B} 4 3 2 (6×1)	ns / ns	6.61 / 6.60		
Groß Schwülper	O	M	S	28	20	13	1.5	4 3 2 2 (9×1)	ns / ns	5.21 / 5.92		
Peine	O	H	S	32	20	11	1.8	9^{B} 2 (9×1)	** / **	6.00 / 7.35		
Celle	O	L	S	36	20	3	6.7	18^{F} 1 1	** / **	0.60 / 4.00		
Han.Münden	W	L	S	12	12	10	1.2	2 2 (8×1)	ns / ns	6.20 / 6.64		
Hes.Oldendorf	W	M	S	42	20	8	2.5	13^{S42} (7×1)	** / **	5.06 / 7.32		
Ströhen	W	H	S	15	15	3	5.0	13^{S42} 1 1	** / **	1.50 / 4.67		
Allerbruch	W	H	S	0	0	0	n.a.	n.a	n.a.	n.a.		
Drakenburg	W	H	S	4	4	3	1.3	2 1 1	n.a.	5.00 / 6.00		
Langwedel	UW	L	S	18	16	8	2.0	8^{B} 2 (6×1)	** / **	5.46 / 7.36		
Brake-Weser	UW	H	S	50	20	7	2.9	11^{B} 3 2 (4×1)	ns / ns	4.42 / 7.43		
Brake-Sieltief	UW	H	S	56	20	8	2.5	6^{B} 5 3 2 (4×1)	ns / ns	6.87 / 7.50		
All strains			W	190	184	65	2.8	21^{B} 17^{sol} 14^{B} 10^{sol} 9 8 7 6 (3×5, 2×4, 4×3, 9×2, 39×1)	ns / **	6.93 / 6.88	S / W	W / W
			s	328	197	60	3.3	29^{S42} 23^{B} 18^{B} 13^{B} 12 11 6 (4×4, 6×3, 8×2, 35×1)	* / **	6.47 / 6.72		

[1] The order of the stations from top to down is in the direction of the water flow from the spring to the estuary.

[2] R = river system, O = Oker/Aller, W = Weser, UW = Unterweser.
[3] E = ecological classification of the chemical quality (exposure) of the River System; L = low, M = moderate, H = high. For details please see the supplementary material
[4] designation of the sampling season; W = winter, S = summer.
[5] as = number of strains which have been isolated from 1 litre of environmental water, s = number of strains that were subjected to ArrayTube hybridization (in winter, most isolates per station where subjected to genetic characterization via ArrayTube hybridization, in summer, however, when more than 20 strains per station were isolated, a random choice of 20 strains were subjected to further genetic characterization), cl = the number of different clones among s strains, s/cl = the number of analysed strains (s) per clone (cl).
[6] The numbers indicate the abundancies of the observed clones in the order of their frequency. For example, in the station Groß Schwülper in winter, 15 clones were found. The most frequent one is represented by 17 strains, the second frequent clone by 8 strains, and 11 clones were represented by only one isolate. The affiliation of the most abundant clones to the exdt_cc groups is indicated by the superscript group name. In the case that the respective clone could not be affiliated to any group (i.e., the clone is so diverged that it has at best only triple-locus-variants in the dataset) the superscript ^{sol} (solitary) is given.
[7] EWS = Ewens-Watterson Slatkin Test for deviation from neutral expectation of the clone-frequency distribution according to the infinite allele model. EW = significant deviation according to the Watterson's homozygosity test, S = significant deviation according to the Slatkin's exact test. The significance codes are
* $0.975 \leq p \leq 0.99$, ** $p > 0.99$; two-tailed test, ns = not significant.
[8] Mean number of pairwise differences (theta π); (s/cl) = as defined above
[9] S and W indicate higher amounts of strains (as) or a higher haplotype ratio (hr) in either summer or winter, respectively. 0 = no difference between summer and winter, n.a. = not applicable.
[10] S and W indicate a higher theta value (theta π; the mean number of pairwise difference) among all strains subjected to ArrayTube hybridization (s) or clones (cl) in either summer or winter, respectively. 0 = no difference between summer and winter, n.a. = not applicable.

Interestingly, the increase of strains in summer does not lead to an increase in genetic diversity, as there is no difference between summer and winter in the genetic diversity of sampling stations when only clones are taken into account (p = 0.19, MWU; column π "(s/cl)", Tab. 1). However, when all strains are taken into account, the genetic diversity per sampling station is significantly lower in summer (p = 0.012, MWU; column π "(s/cl)", Tab. 1). This suggests that the entire genetic diversity is already present in winter and does not increase in summer. Moreover, the increase in strain numbers in summer does not occur uniformly for each clone. It appears as if some clones multiply far stronger than other clones. This is suggested by a substantially larger amount of sampling stations which diverge in summer from the neutral expectation according to the infinite-alleles model (Ewens-Watterson Slatkin test, column "neutrality test", Tab. 1). In these sampling stations, one or few clones are represented by significantly larger numbers of isolates than expected by a neutral distribution of strain numbers. It appears as if the most successful environmental clones belong the ecc group B (see below and see also Figure 1 and 2).

We sought to identify by an AMOVA analysis the parameters which might influence the distribution of the genetic diversity of the *P. aeruginosa* clones, irrespective of their individual strain abundancies. An AMOVA analysis allows user-defined subdivisions of the dataset in order to determine which type of subdivision contributes measurable amounts of variation to the entire dataset. Neither the season (winter vs. summer) nor the geographical connection of the sampling stations via the same river system added any variation to the dataset (Table 2). However, subdivision according to sampling sites of different water quality (see Appendix for Supplemental Table 1) added at least minor variation to the dataset. When taking into account only the low versus strongly polluted sampling sites, the contributed amount of variation (1.34%) is significant (p = 0.028) (Table 2). This suggests to some extend an influence of the sampling site ecology on the distribution of the genetic diversity, which will be analysed more deeply with Fig. 3. Interestingly, the majority of variation (> 90%) is contributed at the level of individual sampling stations (AMOVA, data not shown). This supports the observation that the genetic diversity per sampling stations, taking into account only one strain per clone, is in average the same as in the entire set of environmental strains (column "π (s/cl)"), Tab. 1).

Table 2. AMOVA analysis on the genetic diversity of river isolates of *Pseudomonas aeruginosa*.

Source of variation[a]	% of variation	*P* value
Seasonal (winter vs. summer)	-0.06	0.51337
Ecological (low vs. moderate vs. strongly polluted)	1.2	0.06752
Ecological (only low vs. strongly polluted)	1.34	0.02802
River system (Oker-Aller vs. Weser vs. Unterweser)	0.17	0.44485

[a] The individual sampling places were taken as 'populations' (as termed in the AMOVA nomenclature) (Excoffier *et al.*, 2007) and were affiliated as such into different groups as specified in the first column.

4.2. The extended clonal complexes (ecc) show different population genetic characteristics.

An eBURST analysis revealed the presence of six clonal complexes (cc) (Appendix Figure 11). A cc is defined by a founder clone and all further clones that are related to the founder clone by a series of single locus variants (SLV) (shaded in grey, Fig. 1a). SLVs differ in one out of 17 core genomic markers.

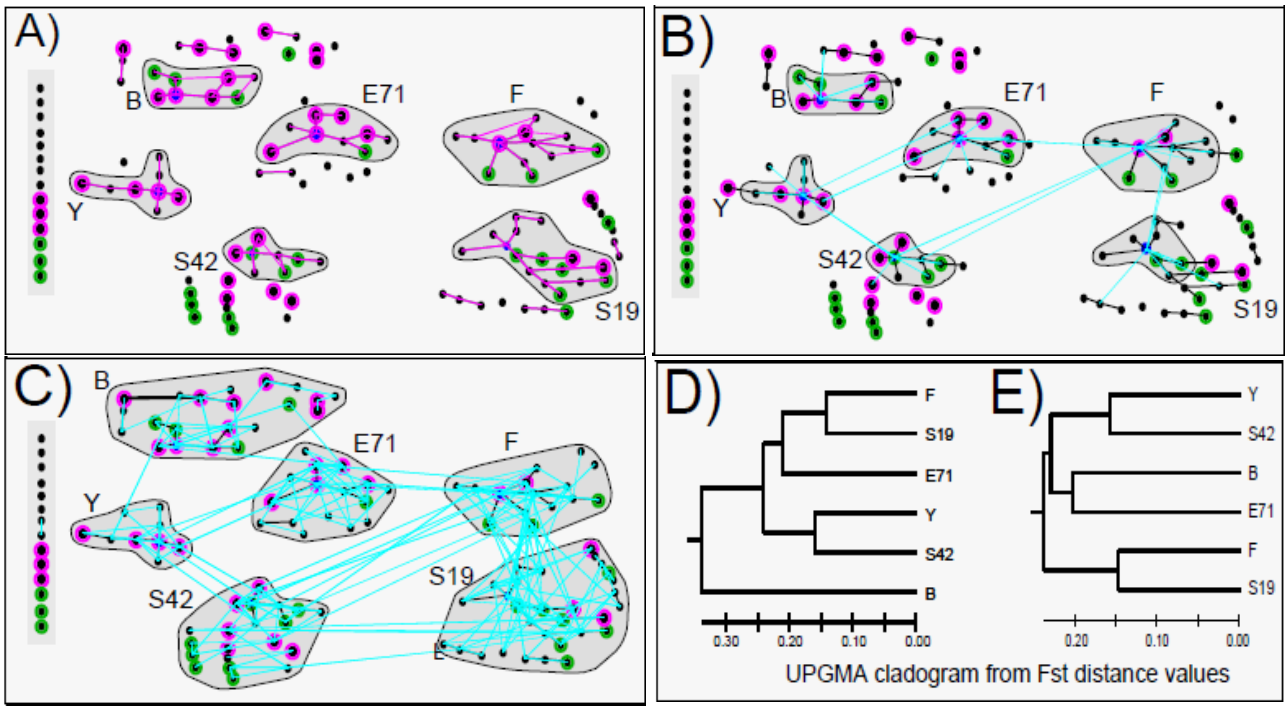

Figure 1. The structure of the genetic relationships among the *P. aeruginosa* host strains as based on eBURST analysis of the core genome SNP pattern. Pink clones are represented by both clinical and environmental strains. Uniquely clinical or environmental clones are coloured in green or in black, respectively. The founder strain of the clonal complexes is coloured in blue. The clones listed in the rectangular grey shaded box left-sided in Fig. A), B) and C) do not have any SLV or DLV within the entire

dataset. (A) The clonal complexes and all their within SLV variants are shaded in grey. The clonal complexes are named according to the clone designation of their founder strains. (B) The SLV and DLV of all founder strains are shaded in grey. Dark grey lines indicate DLV of the founder strains, even if they may belong to other clonal complexes. (C) The clonal complexes and all strains affiliated via a series of DLV with the respective clonal complexes are shaded in grey. All DLV connections present in the dataset are shown by blue lines (both, DLV within an extended clonal complex and DLV across extended clonal complexes). D+E) UPGMA cladograms as deduced from Fst values from the groups as defined in Fig. 2B and 2C, respectively.

A founder is defined to be that clone which has the maximum number of different SLV clones. However, a cc is a biased way to classify clones into genetic groups, as, eg., five locus variants (which differ at five positions from the founder) may be included into a cc as long there is a linking series of SLV clones to the founder (see cc S19, Fig. 1a), whereas double locus variants (DLV) might be excluded if there is no connecting SLV clone in the database (see cc Y, Fig. 1a and b). We therefore sought to classify the clones related to a founder strain in a more unbiased way and at two levels of resolution. The first level encompasses all the SLV and DLV from a given founder, even if a DLV from one founder strain might be member of the cc from another founder strain (Fig. 1b). The second level of resolution encompasses the original cc (Fig. 1a) and all those clones which can be affiliated by DLV connections to one of the clones within a cc. We term this group extended clonal complex (ecc) (Fig. 1c). We realized that such a way of clustering is more reliable than by a phylogenetic tree based on neighbour joining or maximum parsimony methods (data not shown). For the so defined groups at both levels of resolution (Fig. 1b and c) we determined from the clone information (i.e., one strain per clone) the F_{ST} values, a measure for the short term genetic distances between populations (Reynolds et al. 1983; Slatkin, 1995), as input for an UPGMA analysis of genetic similarity of the ecc to each other (Fig. 1d and e), respectively. Both levels of resolution indicate F and S19 as well as Y and S42 to be closely related groups of clones (Fig. 1d and e). At the first level of resolution (Fig. 1b) it appears as if group B is somewhat quite distantly related to all other groups (Fig. 1d). This strong genetic distance is attenuated when also those strains are taken into account, which are more distantly related to the founder strains (Fig. 1c and e).

Group B is obviously the most dominant group, followed by S19, F, and S42 (Table 3). The number of strains per clone, however, deviates substantially among the ecc groups (Table 3).

Table 3: Genetic characteristics of the ecc groups

Statistics	B	E71	Y	S42	F	S19	ALL
Diversity:							
Number of strains (% of all [309])	127 (41.1)	15 (4.9)	15 (4.9)	47 (15.2)	49 (15.9)	56 (18.1)	309 (100)
Number of different strains (% of all haplotypes [88])	17 (19.3)	15 (17.0)	9 (10.2)	12 (13.6)	12 (13.6)	23 (26.1)	88 (100)
Mean number of strains per haplotype	7.47	1	1.66	3.91	4.08	2.43	3.51
Mean number of pairwise differences (all strains)	4.26	3.18	2.06	2.17	2.09	3.92	6.61
Mean number of pairwise differences (only haplotypes)	4.88	3.18	2.44	5.03	2.88	3.91	6.55
Ewens-Watterson neutrality test:							
haplotype configuration	44 19 16 15 7 (3×3) (3×2) (5×1)	15×1	(3×3) (6×1)	33 3 2 (9×1)	18 12 6 3 2 2 (6×1)	7 6 5 (2×4) (3×3) (3×2) (11×1)	44 33 19 18 16 15 12 (2×7) (3×x6) 5 (3×4) (10×3) (9×2) (52×1)
Observed F value	0.18023	N.A.	0.14667	0.50294	0.21949	0.06633	0.05028
Expected F value	0.16204	N.A.	0.15336	0.17726	0.18016	0.08033	0.02682
Watterson F p-value	0.726	N.A.	0.569	1	0.824	0.16	0.99
Slatkin's exact p-value	0.712	N.A.	0.569	1	0.887	0.258	1
linkage disequilibrium:							
I_A^S	0.0566	-0.0082	-0.0184	0.0667	-0.0004	-0.0077	0.0485
P (H_0: $V_D = V_e$)	< 0.0001	0.79	0.89	0.0014	0.44	0.85	< 0.0001
Mean genetic diversity	0.2872	0.1871	0.1438	0.2959	0.1693	0.2297	0.3856
genetic structure	“clonal”	“recombinogenic“	“recombinogenic“	clonal	“recombinogenic“	“recombinogenic“	clonal

The values for diversity and Ewens-Watterson neutrality test were calculated using Arlequin 3.5 (Excoffier *et al.*, 2007).
The null hypothesis for linkage equilibrium for multilocus data was tested using LIAN3.5 (Haubold & Hudson, 2000).

Although the haplotype configuration of all strains affiliated to any of the ecc groups deviates from neutrality according to the EWS test (Table 3), most individual ecc, excluding S42, do not deviate from neutrality. With respect to recombination of genes as a result of potential horizontal gene transfer, the structure of all clones affiliated to the ecc groups is rather clonal, as the standardized Index of Association I_A^S deviates significantly from zero (Table 3, column "all"). However, at the level of individual ecc, the groups E71, Y, F, and S19 appear to be recombinogenic, whereas groups B and and S42 are rather clonal (Table 3).

4.3. The heterogeneous distribution of the ecc groups in sampling stations of different water quality.

We sought to explore whether the six ecc groups (Fig. 2) are uniformly distributed in the environment or whether they differ in abundance or in distribution across sampling places of different water quality (low, moderate, and strongly polluted water).

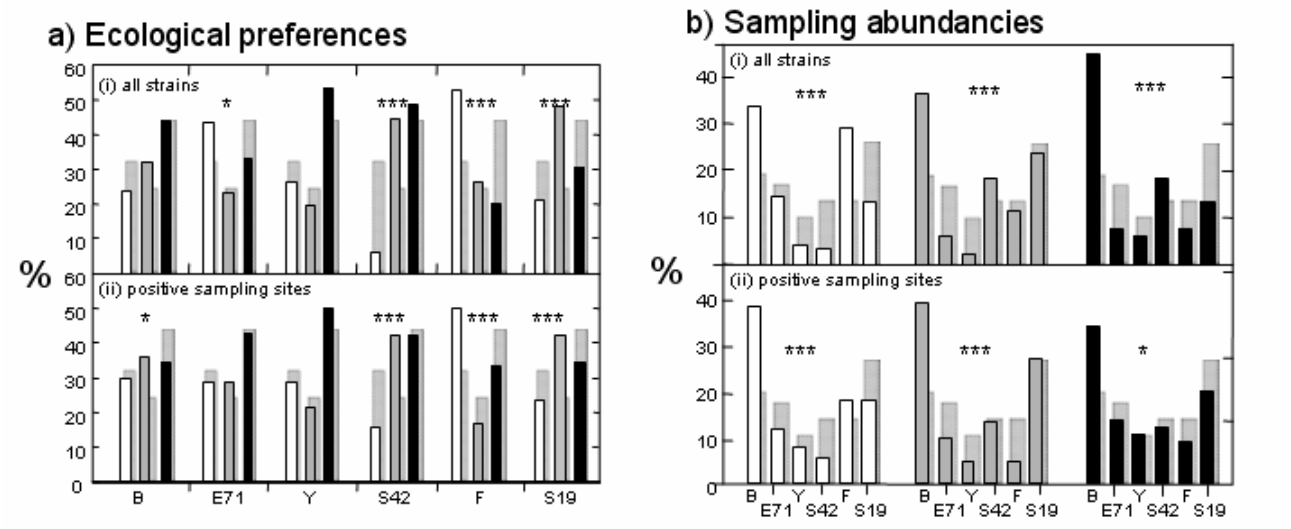

Figure 2. Distribution of *P. aeruginosa* ecc groups in Northern Germany rivers with respect to (a) ecological preferences and (b) sampling frequency. For both (a) ecological preferences and (b) sampling abundancies, the characteristics of the entire dataset is assumed to be the expected distribution under the hypothesis of a uniform distribution of *P. aeruginosa* and is compared to the distributional characteristics of subsets (observed distribution). The observed distributions are shown as white (low), grey (moderate) and black (strong pollution) columns in the foreground, whereas the expected distribution is shown as zigzag shaded column in the background, respectively. The degree of deviation between expected and observed distribution is assessed using the chi-square test. The significance codes for the *p*-values are 0 '***' 0.001 '**' 0.01 '*' 0.05. For both approaches (a) and (b), two measurement criteria are applied, (i) the total amount of isolates, and (ii) the number of successful sampling occasions.

However, with respect to measure (i), the isolation of a large number of strains from, e.g., only a single haplotype of a given ecc group from only a single sampling site per given water quality group (low, moderate or strongly polluted) might erroneously be interpreted as a substantially large environmental distribution and relevance of a given ecc group in a given habitat, though it might just be the result from a very local outburst. This may lead to a bias in data analysis from measure (i). In order to accommodate for such a potential bias, we introduced a second abundance measure (ii), which records only the number of independent successful sampling occasions in which any haplotype of a given ecc group has been isolated from any individual sampling place per given water quality group, irrespective of the number of strains.

For this, we analysed the same dataset following two approaches. First, we focus on the individual ecc group and its distribution across sampling places of different water quality (Fig. 2a). Under the null hypothesis of a uniform distribution we expected a distribution according to the relative proportion of low (32%), moderate (24%), and strongly polluted water (44%) sampling places among all positive sampling places (N=25, see also Tab. 1). Second, we focus on the three different habitat types (low, moderate, and strongly polluted water) and study the relative proportions of the six ecc groups within these habitats. Here, the expected relative proportion is based on the ecc group size, which is the relative proportion of the number of clones per individual ecc group (see also line 2 of Table 3) among all clones encompassed in the six ecc groups (N=88, Table 3) (Fig. 2b).

According to the first approach, group Y is uniformly distributed in the three habitat types (Fig. 3a). In principle, this holds also for groups B and E71, though different types of sampling success measures show slight deviation from a uniform distribution (Fig. 2a). The groups S42, F, and S19, however, differ substantially from the expected distribution, as supported from both measures (i) and (ii) (Fig. 2a). Moreover, these groups differ in the mode of divergence from a uniform distribution. Group S42 is underrepresented in clean water, but rather overrepresented in moderate water. Group F shows a strong overrepresentation in clean water but an underrepresentation in strongly polluted water. Finally, group S19 is underrepresented in clean and strongly polluted, but overrepresented in moderately polluted water.

The second approach (Fig. 2b) focuses on sampling abundancies of the six genetic grouped with respect to the three different habitat types. Interestingly, in all habitat types the actual sampling abundancies differ from the expected sampling abundancies according to the clone sizes of the groups. To a very large extend this is due to the overrepresentation of group B strains in all three habitat types. In contrast, groups E71 and Y appear to be underrepresented in all three habitat types (Fig. 2b). In support of Fig. 2a, also the sampling abundancies depict the underrepresentation of S42 in low polluted

water, the preference of group F for low polluted water, and the underrepresentation of group S19 for both low and strongly polluted water (Fig. 2b).

4.4. Different laboratory survival rates of *P. aeruginosa* in water of different quality.

The results from Fig. 3 suggest that individual ecc groups may show a tendency to prefer or to avoid specific levels of water quality or to show a general overrepresentation irrespective of the water quality. For example, group S42 is hypothesized to avoid clean water (Fig. 3 A and B). In order to test this hypothesis we chose a random subset of clones from S42 and B (five and seven clones, respectively, with one strain per clone) in order to monitor their cell survival in water. In autoclaved distilled water, viable cells of all strains are measurable until day 7. On day 11, however, only one out of five strains (20%) from group S42 still shows viable cells, whereas on day 14 no viable cells from any of the S42 strains were detectable anymore. In contrast, all seven strains (100%) from group B show viable cells until day 25 (Fig. 3 A1). Although, on a qualitative level, all strains from groups S42 and B showed viable cells until day 7, we observed on a quantitative level of the number of viable cells already a substantial decline of S42 in comparison to B. On both day 4 and 7, the strains of group S42 show a lower amount of living cells than the strains from group B (Fig. 3 A2). In sum, these survival experiments on the artificial laboratory environment support the hypothesis from the abundance distribution in the natural environment (Fig. 3).

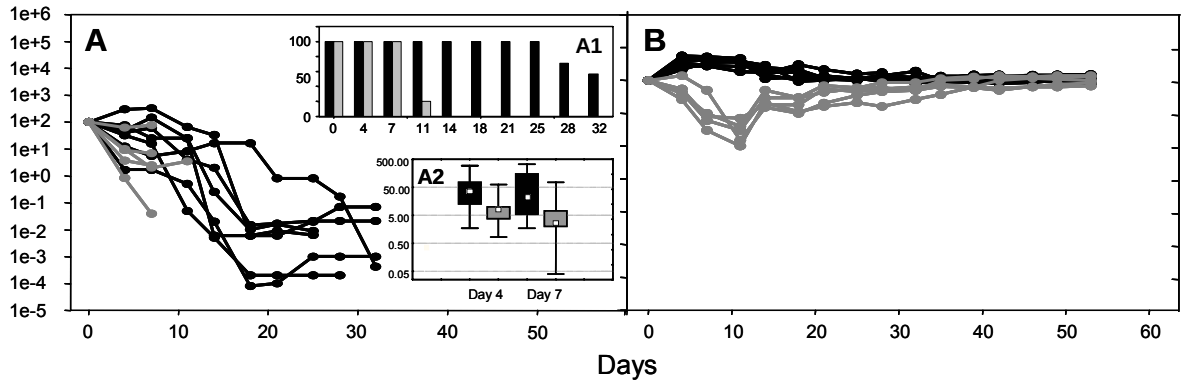

Figure 3. Culturability of strains from the groups B and S42 in distilled water (A) and strongly polluted autoclaved river water (B). Black and grey fills denote group B (seven clones) and group S42 (five clones), respectively. Each clone is represented by a single strain. The y-axis denotes on a logarithmic scale the amount of cells (%). Fig. (A) indicates the amount of culturable cells in comparison to the inoculated number of cells (10^7 cells, 100%). Fig (A1) shows the proportion of strains which were still detectable by culturable cells after the days indicated. The viable cell count data seven group B and five group S42 strains are summarized as box-plots (median, 25%-75% quartiles and the minimum and maximum values) – Fig. (A2).

4.5. Environmental ecc groups differ from clinical ecc groups in motility and virulence phenotypes.

In the environmental strains, for all addressed phenotypes the dominant phenotype is “normal”, with in general more than 60%-80% of all clones per ecc (Fig. 4). Other phenotype variants such as “none”, “minimal”, and “hyper” are usually at a level of 10-20%. In clinical strains, the proportions differ. For example, in swarming and twitching motility, the variants “normal” and “minimal” are approximately equally distributed at 40%-50%. In the phenotype swimming there is a partial reduction in the “normal” variant in ecc groups Y and F, leading respectively to an increase in the variant “minial”. In the phenotype of proteolytic acitivity the difference is even more obvious, as as the variants of “minimal” and “none” often exceed the variant “normal”. In sum, the differences between environmental and clinical ecc groups are significant at “normal” for proteolytic activity, swimming, swarming, and twitching motility ($p = 0.006$ to 0.024, MWU) and at “minimal” for proteolytic activity and swimming ($p = 0.037$ and 0.006, respectively; MWU). Environmental and clinical strains do not differ in their pyocyanin production. The difference between ecc from either environmental or clinical origin is far less pronounced than between environmental or clinical habitats. Only occasionally some substantial differences are observed. For example, the environmental ecc group B has a rather low proteolytic activity, whereas the clinical ecc group B is characterized by a rather large twitiching motility potential (Fig. 4).

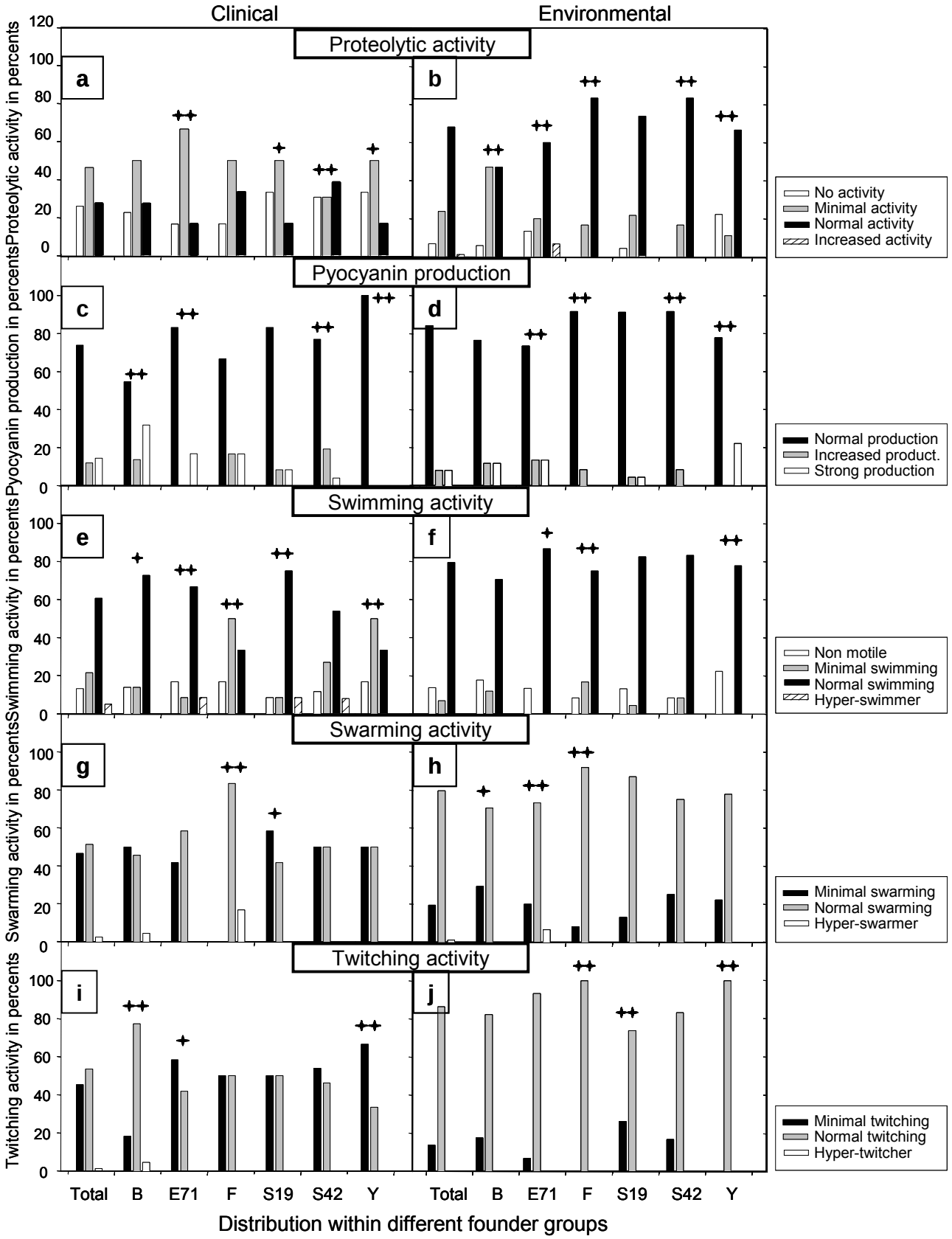

Figure 4. Quantitative determination of virulence and motility phenotypes among the ecc groups from clinical and environmental origin. The semi quantitative category "normal" always indicates the phenotype expressed by the majority of environmental strains. "Total" inidicates all clones that could be affiliated to any of the ecc groups. **p*-value below 0.05, ***p*-value below 0.01 (Chi^2 test).

4.6. Biofilm forming capabilities differ between clinical and environmental strains, but not among the ecc groups.

Within each ecc group, there is no significant difference between environmental and clinical strains when comparing the mean values of the total biovolume (Fig. 5, MWU). Also, across either environmental or clinical origin, there is no significant difference among the ecc groups in their mean values [R package multcomp, (Herberich *et al.*, 2010)]. Nevertheless, the clinical strains obviously show larger extreme values with respect to higher biovolume, which suggests a non-normal distribution and skewness of the clinical data. This is indeed the case, as none of the environmental (excluding B, p = 0.03, Shapiro-Wilk test) but all of the clinical ecc groups (except Y) show significant deviation from normality (p = 0.00001 to 0.0034; Shapiro-Wilk test). Moreover, the clinical ecc groups appear to have also the smaller extreme values than their environmental counterparts, which suggest the clinical groups to have a larger variance. This is the case, as the variance values from the six environmental ecc groups are significantly larger than from the six clinical groups (p = 0.016, MWU). Thus, environmental and clinical strains do not differ in biofilm formation in their mean values, but in a larger spread (variance) around the mean value, with a tendency for few strains per each ecc clinical group to show a substantially increased biofilm formation.

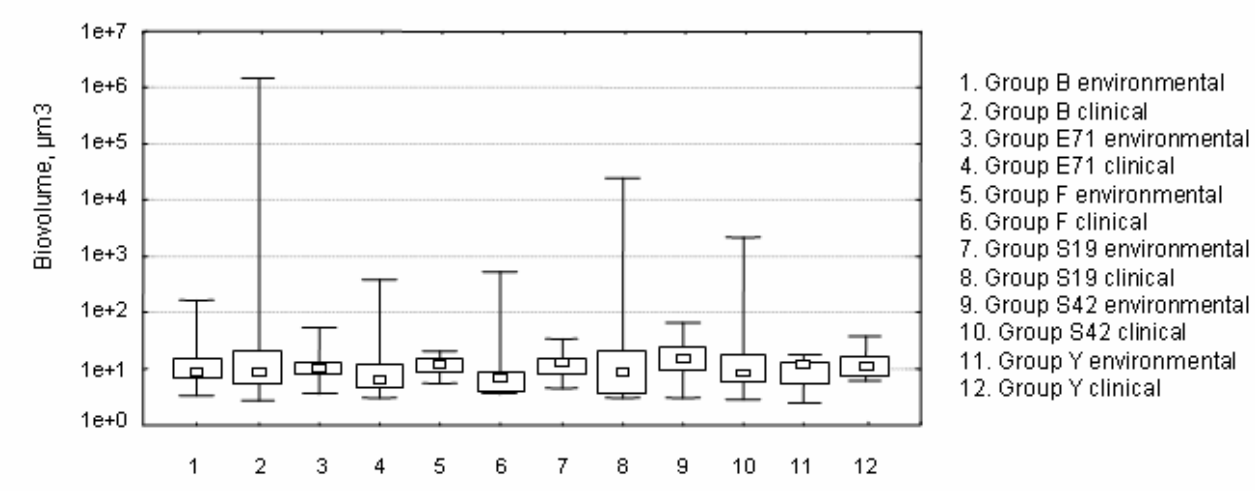

Figure 5. Ability to form biofilm. The box-plots show median, 25-75% quartiles and the minimum and maximum values.

4.7. The phenotypic characterization of seven environmental phages based on their lysis behaviour on 99 environmental clones of *P. aeruginosa.*

Amongst a set of 30 phages isolated from a waste water plant near Braunschweig, Germany (Julia Garbe, PhD thesis, TU Braunschweig; accessible at http://rzbl04.biblio.etc.tu-bs.de:8080/docportal/servlets/MCRFileNodeServlet/DocPortal_derivate_00008650/Doktorarbeit.pdf;jsessionid=DF821E192BF37415A09826A43ADE2AB8), we first explored in detail the lytic behaviour of the most successful seven phages on the set of 99 environmental *P. aeruginosa* clones (one strain per clone).

The proportion of lysed strains (either clear lysis or semi lysis) ranges from approx. 70% (JG005) down to 35% (JG026) (Tab. 4). The relative proportion of clear lysis phenotype among all lysed strains varies considerably from 88% (JG004) down to 12% (JG028) (Tab. 4).

Table 4. The distribution of lysis phenotypes of seven phages across 99 environmental Northern Germany river clones of *P. aeruginosa*

Lysis phenotype	Number of lysed strains by phage						
	JG004	JG005	JG003	JG024	JG025	JG028	JG026
Clear lysis[a]	43 (88)	49 (70)	35 (57)	22 (45)	9 (19)	6 (12)	8 (23)
Semi lysis[a]	6 (12)	21 (30)	26 (43)	27 (55)	39 (81)	46 (88)	27 (77)
No lysis	50	29	38	50	51	47	64

[a] percentage of lysis when the sum of clear lysis and semi lysis is taken as 100%

We sought to determine the phenotypic lysis similarity relationships between the phages. For this we determined for all 21 pairwise combinations of phages the proportion of host strains which are lysed by both phages in the respective combination. For example, from all strains which are lysed by either phage JG025 or JG028, 81% are lysed by both phages, which indicate a high phenotypic similarity between JG025 and JG028 (Appendix Fig. 6a). In contrast, for phages JG004 and JG026, only 27% of strains are lysed by both phages, which indicate a rather low phenotypic similarity. These results suggest two phenotypically distinct groups of phages which we term group *ph*A (phages JG025, JG028, and JG026) and group *ph*B (phages JG004, JG005, and JG003) (Fig. 6). Phage JG024 appears to be phenotypically distinct from both *ph*A and *ph*B (Fig. 6). Group *ph*A appears to be substantially more effective against the hosts strains, as both the ratio of lysis versus non lysis and the ratio of clear lysis versus semi lysis is higher in comparison to group *ph*B (Tab. 4).

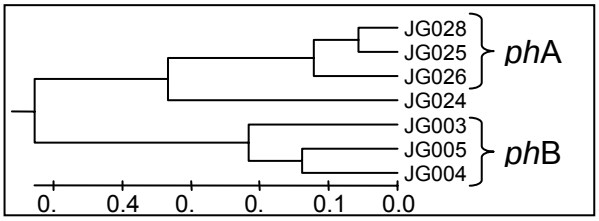

Figure 6. Phenotypic lyses similarity (UPGMA) among seven environmental phages.

4.8. Divergent susceptibility to lytic phage attack among both different ecc groups and among strains from clinical versus environmental origin.

We observed several differences between the six ecc groups in their lysis susceptibility towards both all phages and the individual phage groups phA and phB. Compared across all seven phages, only the clinical exx groups Y and 42 differed signficantly from each other (Fig. 7, Table 5). In the phenotype “no lysis”, the groups B and S42 showed in several instances significantly higher resistance compared to Y and S19 (Fig. 7, Table 5).

Table 5: Significant differences in lysis susceptibility between ecc groups (Fig. 7)

	All seven phages			Phage group phA (JG003, JG004, JG005)			Phage group phB (JG025, JG026, JG028)		
	Complete lysis	Semi lysis	No lysis	Complete lysis	Semi lysis	No lysis	Complete lysis	Semi lysis	No lysis
Environmental strains	n.s.	n.s.	S19 – **B** ** Y – **B** ** Y – **S42** ** S19 – **S42** **	B – **S19** *** B – **S42** ** B – **Y** **	n.s.	n.s.	**S42** – E71*	n.s.	Y – **S42** * Y – **B** *
Clinical strains	**Y** – S42 *	n.s.	Y – **S42** *	B – **E71** * B – **F** *	n.s.	**S42** – S19 *	**Y** – F * **Y** – S19 * **Y** – S42 *	Y – **E71** ** Y – **S42** *	Y – **S42** *
Environmental and clinical strains	n.s.	n.s.	Y – **B** * Y – **S42** * S19 – **S42** *	B – **S19** *	n.s.	n.s.	**Y** – E71 * **Y** – S19 *** **Y** – S42 * **Y** – F *	n.s.	Y – **S42** *

All comparisons were performed using the multcomp package in the open-source software R (http://www.r-project.org/) (Herberich *et al.*, 2010). Only significantly different pairwise combinations are listed. Ecc groups listed in bold and underscore have the larger percentage values as shown in Figure 7. The significance codes are 0 '***' 0.001 '**' 0.01 '*' 0.05 '. n.s. = there is no significant difference among any of the groups.

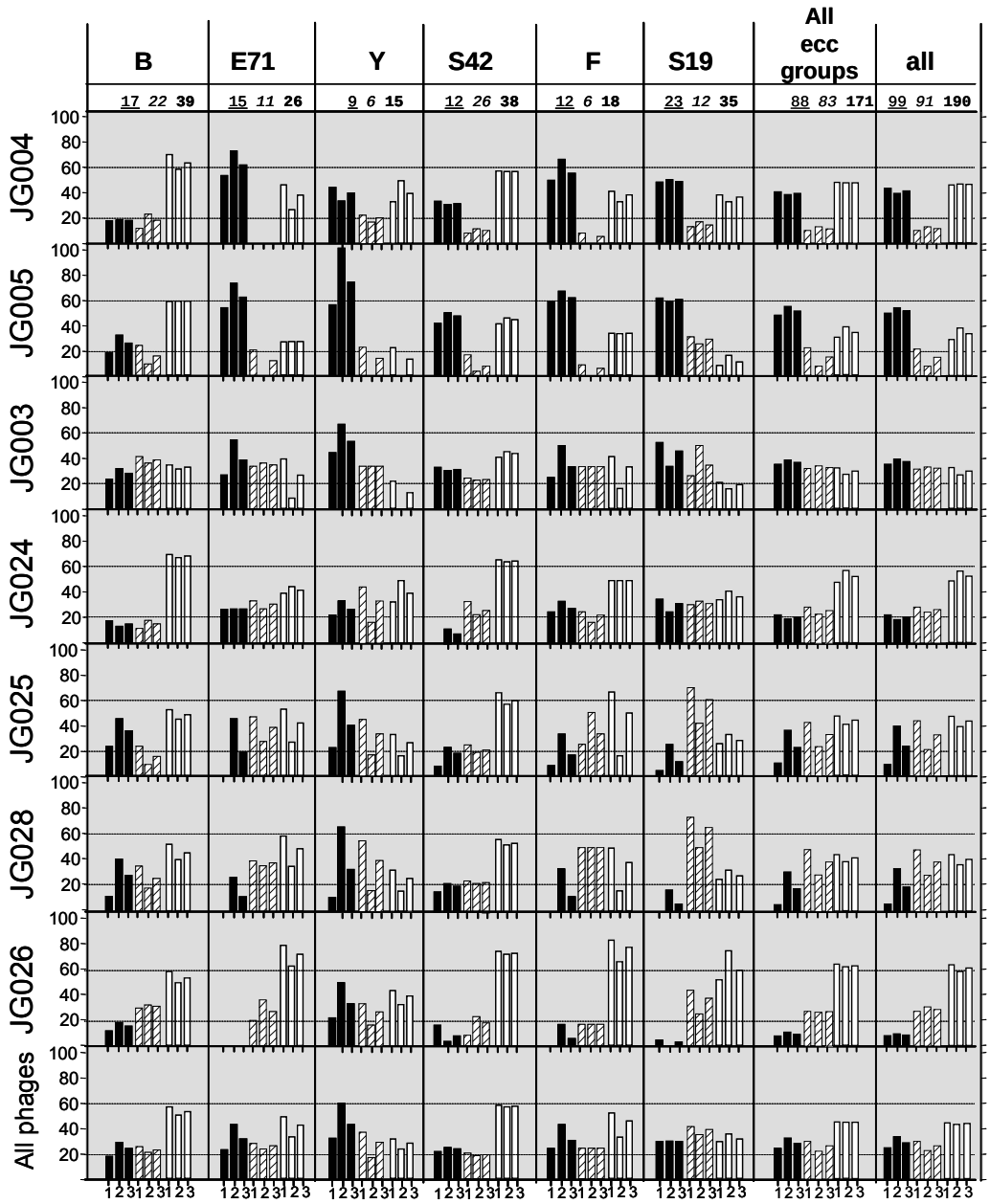

Figure 7. Phage lysis phenotype versus genetic diversity of host strains. The proportions (%) of complete lysis (black bars), semi lysis (shaded bars) and no lysis (white bars) of the individual phages and the average across all phages are plotted against the different genetic groups within the hosts' strains. The column "all ecc" encompasses all strains which could be grouped to any ecc group (see also Fig. 2c). The column "all" encompasses also those strains which do not have any double-locus variant (DLV) within the dataset. Each of the eight subcolumns are headed by a series of three numbers, indicating the amounts of environmental strains (normal underscore), *clinical strains (italic)* and **the sum of both (bold faced).** Within each lysis phenotype, the first bar indicates environmental strains, the second bar indicates clinical strains, and the third bar indicates the sum of both (see the numbering on the x-axis).

At the level of the individual phage groups *ph*A and *ph*B, these differ obviously in their lytic potential towards individual ecc groups. Group B is significantly more resistant to complete lysis by phages from *ph*A than several of the other ecc groups (Fig. 7, Table 5). However, ecc group B does not differ from the other groups with respect to phage group *ph*B. Here, group Y is significantly more sensitive to *ph*B attack (complete lysis) than the other ecc groups. Interestingly, Y differs also in semi-lysis and no lysis by *ph*B from its closely related sister group S42 (Fig. 7, Table 5).

We further observed clinical strains to be apparently more sensitive to phage attack than environmental strains. First, when measured across all seven phages, the ecc groups B and Y showed significantly higher complete lysis rates among the clinical versus environmental strains (p = 0.045 and p = 0.024, respectively; MWU; Fig. 7). Second, when measured across phages JG025 and JG028 only, the clinical strains showed a higher complete lysis ratio among all 6 ecc groups (p = 0.0064 and p = 0.0050, respectively; MWU; Fig. 7). As these two snapshots could suggest a general trend of a larger phage sensitivity of clinical strains, we sought to verify this pattern accross all 42 combinations of phage versus ecc group (7 phages x 6 ecc groups = 42). Interestingly, a significantly larger number of combinations showed a higher proportion of complete lysis phenotype within the clinical strains (30 out of 39, 3 combinations were equal; p = 0.0011, sign-test). Also, in a significantly larger number of combinations the environmental strains showed a higher proportion of the“no lysis” phenotype (27 out of 37, 5 combinations were equal; p = 0.0076, sign-test). This suggests a strong and significant tendency of clinical strains to become more phage sensitive, irrespective of the type of phage and the ecc group affiliation of the host strains.

4.9. The ecc groups differ in their serotype composition.

The most frequent serotypes are 6 and 1, followed by serotypes 3 and 4 (Fig. 8c). In general, this distribution is true for both environmental and clinical strains (Fig. 8a and b).

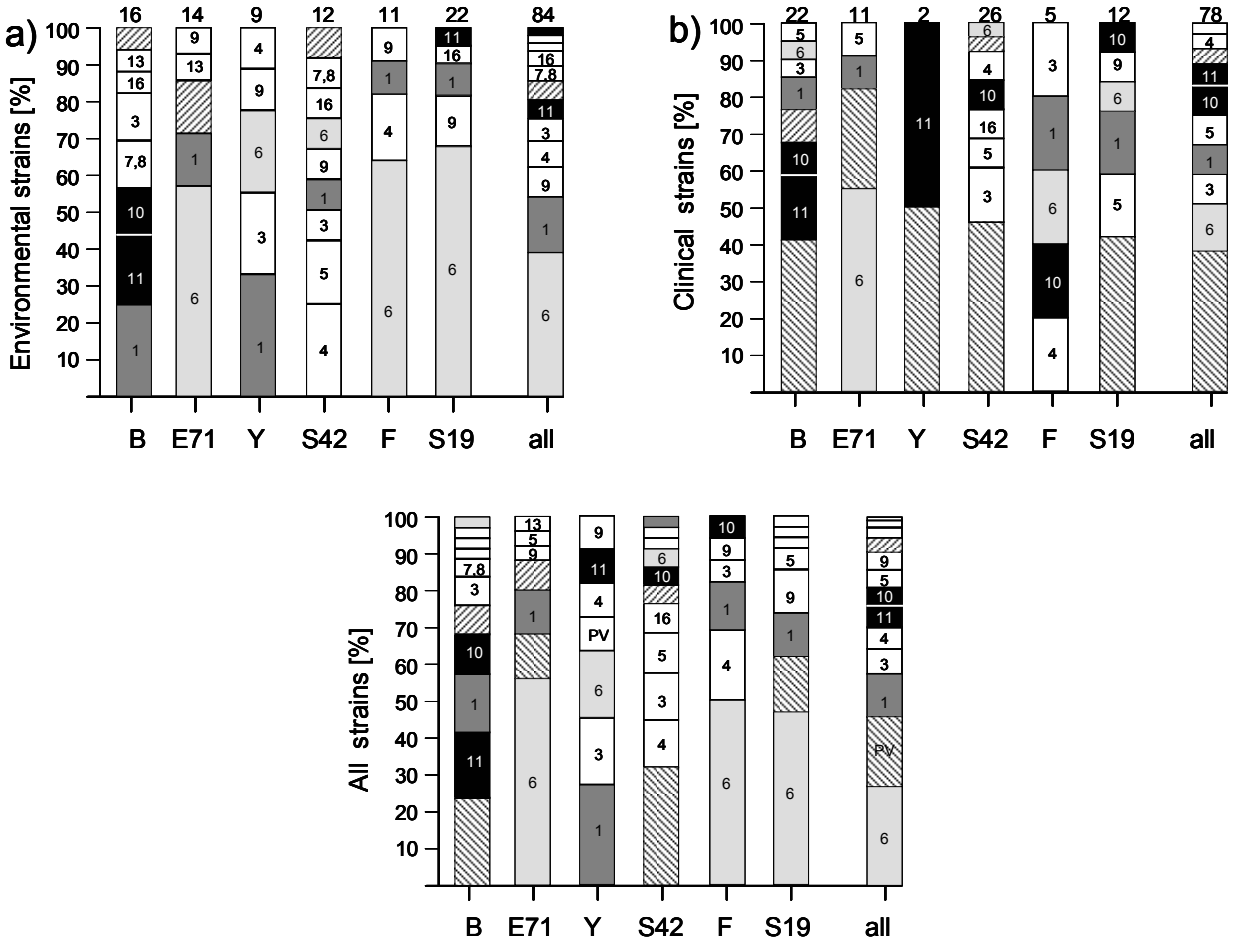

Figure 8. Distribution of the serotypes among the ecc groups. The stacked bars indicate the relative proportions [%] of the serotypes among the six ecc groups from environmental (a) or clinical (b) origin or within the pooled environmental and clinical data (c). The stacked bar designated “all“ encompasses all strains which could be grouped to any ecc group (see also Fig. 1c). The numbers within the bars indicate the serotypes. The designation “7,8“ indicates all serotypes that are either 7 or 8, as these serotypes are often indistinguishable. Stacks without a label indicate serotypes with 3% or less frequency in the respective dataset. The relevant serotypes are coloured in light grey, dark grey, or black. Striped stacks from lower-left to upper-right indicate not typable serogroups; striped bars from upper-left to lower-right indicate polyvalent serotypes. The numbers on top of the bars indicate the absolute numbers of strains per bar.

The major difference between environmental and clinical strains lies in the large proportion of polyvalent serotypes in the clinical strains (Fig. 8b). Interestingly, the

serotypes are far from being uniformly distributed among the ecc groups. E71 is dominated by serotypes 6 (>50%) and 1 (>10%) in both environmental and clinical strains. In contrast, in group B serotype 6 is present only in a small amount (5%) among the clinical strains and not at all in the environmental strains. Instead, group B appears to be dominated by serotypes 10, 11, and 1. The closely related groups Y and S42 (see also Fig. 1d and e) are characterized by a rather uniform distribution of a large number of serotypes, with some small preference for serotypes 1, 3, and 4, but less for serotype 6. In contrast, the closely related groups F and S19 (Fig. 1d and e) show both again a rather large abundance of serotype 6 and 1, and also of serotypes 4 and 9 in the groups F and S19, respectively. In sum, we observe a substantial diversity of serotype distribution among the different ecc groups.

The full table of the phenotypic characteristics and test results is available in Appendix as Supplemental Table 3.

5. Discussion

P. aeruginosa has gained scientific importance because of several reasons. First, it is an opportunistic pathogen causing severe diseases in humans. Second, the genome of *P. aeruginosa* is well equipped with a large number of genes involved in transport and efflux of organic compounds, in intrinsic drug resistance and also in biofilm formation and quorum sensing (Govan & Deretic, 1996; O'Toole & Kolter, 1998a; Poole, 2001; Schuster *et al.*, 2003; Stover *et al.*, 2000b). These features enable *P. aeruginosa* to thrive in a large number of substantially different and also extreme environments. Besides rather common habitats such as soil, water, plants, and different animal and human body compartments, *P. aeruginosa* replicates also in sodium dodecyl sulfate (Hagelueken *et al.*, 2006) and contrast media (Tress *et al.*, 1994). This impressive versatility, together with its ubiquitous distribution, has lead to the notion that *P. aeruginosa* strains thrive equally well in different habitats. As such, it has been frequently stated that there is no difference between *P. aeruginosa* strains from different environmental or clinical habitats (Alonso *et al.*, 1999; Foght *et al.*, 1996; Kiewitz & Tummler, 2000; Morales *et al.*, 2004; Ruimy *et al.*, 2001).

A consistent interpretation of this notion would imply that *P. aeruginosa* is hardly affected by selective pressures, as its genetic versatility may readily repel any biotic or abiotic threats. If this were true, in consequence, natural genetic drift would be responsible as major factor of any observable population structure in *P. aeruginosa*. Although natural genetic drift is unquestionably an important evolutionary force (Hartl & and Clark, 2007; Lynch & Conery, 2003; Lynch, 2006), drift has been found to be rather dominant in bacteria with a very narrow ecological niche such as obligate endosymbionts (Herbeck *et al.*, 2003; Kuo *et al.*, 2009), whereas *P. aeruginosa* is ecologically ubiquitously distributed. Indeed, numerous studies show the susceptibility of *P. aeruginosa* to selective pressures while microevolutionary adapting to the human habitat (Ciofu *et al.*, 2010; Hogardt *et al.*, 2007; Smith *et al.*, 2006a; Sriramulu *et al.*, 2005). Therefore, in contrast to indirect implications (Alonso *et al.*, 1999; Foght *et al.*, 1996; Kiewitz & Tummler, 2000; Morales *et al.*, 2004; Ruimy *et al.*, 2001), it is reasonable to expect that selective pressure exerted by different types of natural environments may also affect *P. aeruginosa* in the natural environment.

The contemporary consensus is that *P. aeruginosa* is characterized by a non-clonal epidemic population structure. “Epidemic” is defined by the world health organisation as “the occurrence in a community or region of cases of an illness, specific health-related behaviour, or other health-related events clearly in excess of

normal expectancy" (http://www.who.int/hac/about/definitions/en/index.html status 22 June 2010). Adopted to *P. aeruginosa*, "epidemic" apparently refers to certain frequently sampled clones, e.g., the famous clone C (Romling *et al.*, 1994; Romling *et al.*, 1995) or the entirely clinical MDR serotype O:12 group (Giammanco *et al.*, 1985; Pirnay *et al.*, 2009; Watine, 1999). On a geological scale, epidemic clones may be the result of sudden outbursts and may therefore disappear also similarly sudden, once the acute selective advantage diminishes. However, also the remainder and mostly majority of the genetic diversity of a species exhibit usually a characteristic genetic structure, which may be far from a uniform mix. Often, the genetic diversity of a species assembles into more or less distinct genetic groups or lineages, each of which may consist of numerous different clones. In contrast to epidemic clones, which may appear and vanish again (Achtman, 1997), these lineages are rather robust on a geological scale. Numerous studies have shown such lineages also to exist in *P. aeruginosa*, though they have until now received only few attention. In our study, we term these lineages extended clonal complexes (ecc). Here, we study the characteristics of these ecc with respect to environmental and clinical habitats. We will discuss selective forces which appear to be cohesive for the ecc and which thereby may drive the further diversification (cladogenesis) of these lineages, irrespective of occasionally emerging epidemic clones. Unless going extinct, it is reasonable to expect that they will ultimately lead to the emergence of novel species from the current *P. aeruginosa* pool.

5.1. Nonrandom distribution of extended clonal complexes in German river: eccB emerges as a dominant aquatic ecotype

The AMOVA study indicated that neither the season of isolation (winter versus summer) nor the geographical distribution of isolates (river Weser versus river Oker/Aller) appeared to contribute any influence on the genetic diversity of isolates (Table 2). We do, however, observe a larger amount of isolates in summer than in winter, probably as a result of the higher temperature which may enable better growth (Table 1) (Pirnay *et al.*, 2005). Yet, neutrality tests on the haplotype configuration indicate that not all clones appear to benefit equally well from the temperature shift. In contrast, it appears as if only few clones benefit and increase disproportionally high (Table 1), which may suggest that some clones may respond differently to environmental challenges. In accord with this observation, the AMOVA study suggested some influence of the water quality of sampling stations (as defined by the chemical load determined by the Lower Saxony Ministry of the Environment) on the genetic diversity of the strains (Table 2). This is in so far interesting as, compared to the tremendous differences in the habitats in which *P. aeruginosa* may

reside globally, the differences in water quality appear to be marginal. Motivated from these insights, we sought to explore if potential differences in response to environmental challenges relate to the genetic lineages (extended clonal complexes, Fig. 1) we identified. Indeed, some ecc apparently prefer rather clean water compared to strongly polluted water (eccF), others thrive rather in strongly polluted water and appear to avoid clean water (eccS42), whereas others again are equally abundant in all three levels of water quality (eccB, E71) (Fig. 2). Survival experiments in laboratory assays (Fig. 3) supported this observation and suggested rather direct and immediate effects on the viability/culturability of the cells.

Another interesting feature is the dominance of eccB in our samples, irrespective of the water quality (Fig. 2). Specifically in summer, the eccB group is often the most sampled group per sampling station. From a genetic perspective, eccB is quite distinct from the other ecc, especially when taking only the founders and its closest relatives into account. These findings suggest eccB to be better adapted to environmental water than the remainder of *P. aeruginosa*, i.e., to constitute a potential water ecotype. We do note that we base our assumptions only on cultivation data, whereas molecular methods might show alternative results. Even if this were the case, the readily culturability of eccB may also suggest a better physiological adaptation to environmental water. For a global comparison, we have kindly been supplied from different authors with representative strains from previous studies (see Appendix for Supplemental Table 2). Interestingly, our findings of eccB to be a perhaps specific water ecotype gains support from a recent study on water isolates in Japan. Khan et al identified eccB to be distinct open ocean isolates which were rather absent from fresh water river or bay water (Khan *et al.*, 2007). The authors supported their hypothesis of a better adaptation of the open ocean isolates to higher salinity by laboratory microcosm survival experiments, in which the marine strain performed at 7% NaCl indeed better than the freshwater or clinical isolate (Khan *et al.*, 2007). Thus, the study by Khan et al supports the hypothesis of eccB being a probably characteristic environmental water ecotype within *P. aeruginosa*. Also, among 30 isolates from Japanese rivers isolated more than 35 years ago (Mutharia *et al.*, 1982), eccB was the dominant group (together with eccF, each 8 isolates). Another, indirect, evidence comes from a recent very comprehensive study of 328 globally spread isolates, amongst them 55 environmental water isolates (Pirnay *et al.*, 2009). In this study, eccB is equivalent to the MST groups F and G (Pirnay *et al.*, 2009). Comparable to our results and to the japanese water isolates (Khan *et al.*, 2007; Khan *et al.*, 2008), the MST groups F and G are obviously genetically and phenotypically distinct from the rest of *P. aeruginsa* (Pirnay *et al.*, 2009). In the eccB group of that study (Pirnay *et al.*, 2009), the environmental water clones are

significantly overrepresented in comparison to the other groups. In sum, our data and those from several other studies support the notion of eccB to be an evolutionary lineage being better adapted to environmental water than other lineages of *P. aeruginosa*. However, whereas a study on the *P. aeruginosa* diversity of the Belgian river Woluwe also identifies the eccB group to be genetically and phenotypically different from the other *P. aeruginosa* groups, eccB was not abundantly sampled in this study (Pirnay *et al.*, 2005). Thus, in terms of sampling abundancy of eccB, the river Woluwe in Belgium obviously differs from the German river systems of Oker/Aller and Weser.

Besides its genetic distance from other ecc and apparently superior environmental adaptation, eccB differs from other *P. aeruginosa* lineages by its exoS/exoU characteristics. ExoS and ExoU are virulence effector molecules. These are injected into the host cell by the type III secretion system. It has been reported that ExoS is an important toxin required for colonization and dissemination during infection, whereas ExoU is mainly associated with increased virulence (Lin *et al.*, 2006; Schulert *et al.*, 2003; Shaver & Hauser, 2004). In the overwhelming majority of cases (>98%), a strain has either the *exoS* or *exoU* gene, whereas having both or neither genes is extremely seldom. The most common combination of these virulence factors is the combination $exoS^+/exoU^-$, which is found in the literature at a frequency of approximately 70% of all yet studied strains, followed by the second frequent combination of $exoS^-/exoU^+$ (Feltman *et al.*, 2001; Pirnay *et al.*, 2009). In accordance with the previously published results, the combination $exoS^+/exoU^-$ is the most abundant (around 73%) in the here studied population, also, if to consider the population as whole, around 27% of it, has the combination of $exoS^-/exoU^+$, present in all ecc, irrespective of environmental or clinical origin. However, this combination of $exoS^-/exoU^+$, we find in our environmental strains only in eccB (in both environmental and clinical strains at approximately 60%), and to a very low frequency (7%) in the clinical eccS42 strains. The very seldom combination of $exoS^-/exoU^-$ is also present only in group eccB, in the most abundant haplotype B30 (see Figure 12 in the Appendix section). It is currently unclear if this apparent linkage of adaptation to environmental water bodies and the presence of the exoU gene are of any biological relevance or if this is just a result of a random evolutionary chance.

In sum, our findings show that the dispersal of *P. aeruginosa* clones in different environmental water bodies is by far not random or uniform. This indicates that the microevolutionary dynamics of *P. aeruginosa* is substantially affected by the natural environment (Khan *et al.*, 2007).

5.2. Phages as driver of cladogenesis in *P. aeruginosa*

It is well accepted that viral numbers exceed those of bacteria, and that lytic phages as predators are a major source of mortality of bacteria in aquatic habitats (Suttle, 2007; Weinbauer, 2004). This leads to dynamic phage-host relationships, which indicates the huge impact phages may have on the evolution of their hosts (Forterre, 2006).The mortality pressure of phages resulted in the increasing interest of phages as therapeutical agents for microbial infection diseases (Debarbieux *et al.*, 2010; McVay *et al.*, 2007; Merabishvili *et al.*, 2009; Thiel, 2004; Watanabe *et al.*, 2007). Also *P. aeruginosa* is under a strong worldwide selective pressure by phages (Ceyssens *et al.*, 2009). Though phages show frequently a quite narrow host spectrum, limited to one or few species, phages usually do not lyse all strains of a single species (Weinbauer, 2004). Recent rather extensive studies on the phage-host relationships in *P. aeruginosa* documented that even the most effective phages do form plaques on only approximately 55% of the tested host strains (Ceyssens *et al.*, 2009; Knezevic *et al.*, 2009). Interestingly, it has until now not been explored if the phenotype of resistance or susceptibility to phages is rather uniformly distributed across the hosts diversity or not. Such insight, however, would allow addressing the impact of phages on the emergence, cohesion and further divergence of the deeper phylogenetic lineages within a species.

To study this, we applied seven environmental phages which we classified into two similarity groups (*ph*A and *ph*B) each consisting of three phages according to their lysis phenotypes. These phages do differ significantly in their lysis behaviour across the ecc groups. This is not only the case for the sum of all phages, but also for the individual phage groups *ph*A and *ph*B (Table 5). In our quest to get a first insight into the nature of this observation, we determined the serotype composition of the ecc groups. The type of O-antigen lipopolysaccharid structure of cell membranes, which is determinative for the serotype, often also serves as specific receptor for phages Yokota (Yokota *et al.*, 1994). We observed that the ecc groups showed characteristic and different serotype composition (Fig. 8). Particularly eccB is characterized by a rather different serogroups composition compared to the other groups. Additionally, we observed that the type of the serogroup was somewhat predictive for the lysis success, ranging from being susceptible to all phages (serotype 3) to providing resistance to phage attack (serotype 4, 7, 8) (see Supplemental material, Fig. 9). The consolidation of the results from genetic structure (Fig. 1), serotype composition of ecc groups (Fig. 8) and the lysis susceptibility characteristics of serotypes (see Supplemental material, Fig. 9) suggest a

mechanism which may explain the differential lysis phenotype of the phages on ecc groups. Though this might be mostly the case for the environmental populations of *P. aeruginosa*, as clinical strains, especially the ones that induce chronic infections, tend to change their lipopolysaccharide profile in the direction of gaining polyagglutination, autoagglutination and “no agglutination” phenomena (Hancock *et al.*, 1983; Pitt *et al.*, 1986). Furthermore, these results suggest phages to contribute by their distinguishing lysis phenotype to the inner cohesion of the ecc and thereby to the further cladogenesis within *P. aeruginosa*.

Another interesting feature is the observation that the clinical strains obviously become more susceptible to phage lysis, irrespective of the individual phages and ecc group. The molecular reason for this is currently unclear. The here observed larger susceptibility of clinical strains nourishes the hope of successful application of phages in phage therapy (Debarbieux *et al.*, 2010; McVay *et al.*, 2007; Merabishvili *et al.*, 2009; Thiel, 2004; Watanabe *et al.*, 2007).

5.3. Natural environment versus clinics: Which habitat has a greater impact on the microevolutionary dynamics in *P. aeruginosa*?

Numerous studies, including ours, have shown that in general representatives from all genetic lineages of *P. aeruginosa* are found both in the environmental and clinical habitat (Alonso *et al.*, 1999; Kiewitz & Tummler, 2000; Morales *et al.*, 2004; Ruimy *et al.*, 2001). There may be very few exceptions, such as the MDR serotype 0:12 lineage, which has not yet been found in environmental samples. Numerous studies have also shown that *P. aeruginosa* strains may substantially change genetically and phenotypically the longer they reside in the clinic habitat (which is set here equivalent to animal or human bodies). The dominance of studies with a clinical motivation reflects our interest in *P. aeruginosa* due to its threat as severe opportunistic human pathogen. Nevertheless, leaving our anthropocentric interest in *P. aeruginsa* aside for a while, the question is justified whether the clinical habitat or the natural environment has a larger influence on the evolution of *P. aeruginosa*.

We argue here that the natural environment is probably far more important for the microevolutionary dynamics in *P. aeruginosa* than the clinical habitat. In order to discuss this, it is important to reflect census population sizes, selection pressures and the migrational move between environmental and clinical habitats.

First, it is reasonable to assume that the fraction of *P. aeruginosa* cells residing in the clinical habitat is rather tiny compared to the environment. Only a few individuals from the humans and other animals are being infected, either via environmental or nosocomial sources and probably not all *P. aeruginosa* cells from a dead human

body will return to the environment again. In contrast, it appears as if *P. aeruginosa* is present in all sorts of non-extreme environmental water sources (Ahlen *et al.*, 1998; Khan *et al.*, 2007; Mena & Gerba, 2009; Pellett *et al.*, 1983; Pirnay *et al.*, 2005; Romling *et al.*, 2005). Taking into account the vast amount of environmental water bodies, the census environmental population size is probably far larger than the clinics population size.

Second, in order for the clinical habitat to have a substantial impact on the global population of *P. aeruginosa*, it would first be necessary for the clinical strains to (i) return back into the environmental habitat again, and to (ii) then successfully replace native environmental strains. For the latter, this would necessitate clinical strains to gain, via evolutionary change in the infected animal body, those phenotypes which might then confer a selective advantage in the environmental habitats. We have studied here several phenotypes in comparison of clinical and environmental strains. As a general observation, clinical strains tend to loose or weaken environmentally relevant abilities as a result of the clinical habitat, such as proteolytic activity and several motility phenotypes. Other phenotypes, such as pyocyanin production, did not change. Interestingly, we observed for a subset of clinical strains a tendency to perform larger biofilms. Biofilms are since long known to be a defence mechanism, as biofilms provide an orders of magnitude larger resistance to antiobiotics than in planctonic cells (Lewis *et al.*, 2005; Lewis, 2001). The potential enlargement of biofilm formation could be of adaptive advantage in the environment. However, in some instances, biofilms may also readily be destroyed by phages (Hughes *et al.*, 1998; Knezevic & Petrovic, 2008; Lewis, 2001; Suthereland *et al.*, 2004), therefore an increased biofilm formation ability might not necessarily convey a selective advantage. Beyond the results presented here, as has nicely been compiled by Ciofu et al (Ciofu *et al.*, 2010), further mutational changes in clinical isolates (mostly CF) are increased alginate production and occurrence of mucoid variants (Govan & Nelson, 1993), loss of quorum sensing (D'Argenio *et al.*, 2007; Smith *et al.*, 2006a), loss of motility (Mahenthiralingam *et al.*, 1994), loss of effector proteins of the type III secretion system (Jain *et al.*, 2004), loss of the O-antigen components of the lipopolysaccharide (Hancock *et al.*, 1983), reduced virulence (Luzar & Montie, 1985), reduced capacity for in vitro biofilm formation (Lee *et al*., 2005) (though this contrasts with our observations, Fig. 5) and increased antibiotic resistance (Ciofu *et al.*, 2001). Thus, our own observations and insights from literature indicate that evolutionary changes of *P. aeruginosa* in clinical habitats are mainly associated with losses of phenotypes due to the specialisation to the human habitat. Moreover, our studies indicate that these phenotypic changes affect all ecc groups more or less equal. It can therefore be assumed that the return of clinically adapted *P. aeruginosa* into the

ruggedness of natural habitats is associated with negative selection and therefore a decreased potential to successfully replace the far more abundant native environmental clones.

In sum, we do argue here that the environmental habitat is the major habitat shaping the population structure of *P. aeruginosa* (Khan *et al.*, 2007). The previous mainly clinically based population genetic studies probably simply reflect the genetic structure present, and shaped by, the natural environment. Among the epidemic clones discussed, probably only very few may be indeed epidemic in a clinical sense (see WHO definition). A typical example for this might be the MDR serotype O12 clone (Patzer & Dzierzanowska, 1994; Talarmin *et al.*, 1996; Watine, 1999). The relative abundance of clone C in the clinical habitat, however, may simply reflect the observation that clone C is also a very abundant clone in aquatic habitats, but may perhaps not be a typically epidemic clone.

We therefore suppose that *P. aeruginosa* mainly evolves in the natural environment and that clinically triggered evolutionary changes play only a minor role in the global evolution of *P. aeruginosa*.

6. References:

Achtman, M. (1997). Microevolution and epidemic spread of serogroup A *Neisseria meningitidis*--a review. *Gene* **192**, 135-140.

Achtman, M., Morelli, G., Zhu, P. & other authors (2004). Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc Natl Acad Sci U S A* **101**, 17837-17842.

Ahlen, C., Mandal, L. H. & Iversen, O. J. (1998). Identification of infectious *Pseudomonas aeruginosa* strains in an occupational saturation diving environment. *Occup Environ Med* **55**, 480-484.

Ajayi, T., Allmond, L. R., Sawa, T. & Wiener-Kronish, J. P. (2003). Single-Nucleotide-Polymorphism Mapping of the *Pseudomonas aeruginosa* Type III Secretion Toxins for Development of a Diagnostic Multiplex PCR system. *J Clin Microbiol* **41**, 3526-3531.

Alonso, A., Rojo, F. & Martinez, J. L. (1999). Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin. *Environ Microbiol* **1**, 421-430.

Altoparlak, U., Aktas, F., Celebi, D., Ozkurt, Z. & Akcay, M. N. (2005). Prevalence of metallo-beta-lactamase among *Pseudomonas aeruginosa* and *Acinetobacter baumannii* isolated from burn wounds and in vitro activities of antibiotic combinations against these isolates. *Burns* **31**, 707-710.

Anjaiah, V., Cornelis, P. & Koedam, N. (2003). Effect of genotype and root colonization in biological control of fusarium wilts in pigeonpea and chickpea by *Pseudomonas aeruginosa* PNA1. *Can J Microbiol* **49**, 85-91.

Auch, A. F., von Jan, M., Klenk, H.-P. & and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* **2**, 117-134.

Bagge, N., Hentzer, M., Andersen, J. B., Ciofu, O., Givskov, M. & Hoiby, N. (2004a). Dynamics and spatial distribution of beta-lactamase expression in *Pseudomonas aeruginosa* biofilms. *Antimicrob Agents Chemother* **48**, 1168-1174.

Bagge, N., Schuster, M., Hentzer, M., Ciofu, O., Givskov, M., Greenberg, E. P. & Hoiby, N. (2004b). *Pseudomonas aeruginosa* biofilms exposed to imipenem exhibit changes in global gene expression and beta-lactamase and alginate production. *Antimicrob Agents Chemother* **48**, 1175-1187.

Baker, N. R. (1982). Role of exotoxin A and proteases of *Pseudomonas aeruginosa* in respiratory tract infections. *Can J Microbiol* **28**, 248-255.

Baltch, A. L. a. R. P. S. e. (1994). *Pseudomonas aeruginosa: Infections and Treatment.* New York, NY: Marcel Dekker, Inc.

Bano, N. & Musarrat, J. (2003). Characterization of a new *Pseudomonas aeruginosa* strain NJ-15 as a potential biocontrol agent. *Curr Microbiol* **46**, 324-328.

Bapteste, E., O'Malley, M. A., Beiko, R. G. & other authors (2009). Prokaryotic evolution and the tree of life are two different things. *Biol Direct* **4**, 34.

Baron, S. S. & Rowe, J. J. (1981). Antibiotic action of pyocyanin. *Antimicrob Agents Chemother* **20**, 814-820.

Barth, A. L. & Pitt, T. L. (1995). Auxotrophic variants of *Pseudomonas aeruginosa* are selected from prototrophic wild-type strains in respiratory infections in patients with cystic fibrosis. *J Clin Microbiol* **33**, 37-40.

Beatson, S. A., Whitchurch, C. B., Sargent, J. L., Levesque, R. C. & Mattick, J. S. (2002a). Differential regulation of twitching motility and elastase production by *Vfr* in *Pseudomonas aeruginosa*. *J Bacteriol* **184**, 3605-3613.

Beatson, S. A., Whitchurch, C. B., Semmler, A. B. & Mattick, J. S. (2002b). Quorum sensing is not required for twitching motility in *Pseudomonas aeruginosa*. *J Bacteriol* **184**, 3598-3604.

Behar, A., Yuval, B. & Jurkevitch, E. (2008). Gut bacterial communities in the Mediterranean fruit fly (*Ceratitis capitata*) and their impact on host longevity. *J Insect Physiol* **54**, 1377-1383.

Bicanic, T. A. & Eykyn, S. J. (2002). Hospital-acquired, Native Valve Endocarditis Caused by *Pseudomonas aeruginosa*. *Journal of Infection* **44**, 137-139.

Byrd, J. J., Xu, H. S. & Colwell, R. R. (1991). Viable but nonculturable bacteria in drinking water. *Appl Environ Microbiol* **57**, 875-878.

Ceyssens, P.-J., Noben, J.-P., Ackermann, H.-W. & other authors (2009). Survey of *Pseudomonas aeruginosa* and its phages: *de novo* peptide sequencing as a novel tool to assess the diversity of worldwide collected viruses. *Environmental Microbiology* **11**, 1303-1313.

Chester, I. R., Meadow, P. M. & Pitt, T. L. (1973). The relationship between the O-antigenic lipopolysaccharides and serological specificity in strains of *Pseudomonas aeruginosa* of different O-serotypes. *J Gen Microbiol* **78**, 305-318.

Chitkara, Y. K. & Feierabend, T. C. (1981). Endogenous and exogenous infection with *Pseudomonas aeruginosa* in a burns unit. *Int Surg* **66**, 237-240.

Chugani, S. & Greenberg, E. P. (2007). The influence of human respiratory epithelia on *Pseudomonas aeruginosa* gene expression. *Microb Pathog* **42**, 29-35.

Ciofu, O., Fussing, V., Bagge, N., Koch, C. & Hoiby, N. (2001). Characterization of paired mucoid/non-mucoid *Pseudomonas aeruginosa* isolates from Danish cystic fibrosis patients: antibiotic resistance, beta-lactamase activity and RiboPrinting. *J Antimicrob Chemother* **48**, 391-396.

Ciofu, O., Mandsberg, L. F., Bjarnsholt, T., Wassermann, T. & Hoiby, N. (2010). Genetic adaptation of *Pseudomonas aeruginosa* during chronic lung infection of patients with cystic fibrosis: strong and weak mutators with heterogeneous genetic backgrounds emerge in *mucA* and/or *lasR* mutants. *Microbiology* **156**, 1108-1119.

Claverys, J. P., Prudhomme, M., Mortier-Barriere, I. & Martin, B. (2000). Adaptation to the environment: *Streptococcus pneumoniae*, a paradigm for recombination-mediated genetic plasticity? *Mol Microbiol* **35**, 251-259.

Cohan, F. M. (2001). Bacterial Species and Speciation. *Syst Biol* **50**, 513-524.

Cohan, F. M. (2002). What are becterial species? *Annu Rev Microbiol* **56**, 457-487.

Cohan, F. M. & Perry, E. B. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* **17**, R373-386.

Cohan, F. M. & Koeppel, A. F. (2008). The origins of ecological diversity in prokaryotes. *Curr Biol* **18**, R1024-1034.

Colwell, R. R. (1970). Polyphasic taxonomy of the genus *Vibrio*: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. *J Bacteriol* **104**, 410–433.

Connor, N., Sikorski, J., Rooney, A. P. & other authors (2010). Ecology of speciation in the genus *Bacillus*. *Appl Environ Microbiol* **76**, 1349-1358.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R. & Lappin-Scott, H. M. (1995). Microbial Biofilms. *Annual Review of Microbiology* **49**, 711-745.

Costerton, J. W., Stewart, P. S. & Greenberg, E. P. (1999). Bacterial Biofilms: A Common Cause of Persistent Infections. *Science* **284**, 1318-1322.

Curran, B., Jonas, D., Grundmann, H., Pitt, T. & Dowson, C. G. (2004). Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. *J Clin Microbiol* **42**, 5644-5649.

Curran, B., Morgan, J. A., Honeybourne, D. & Dowson, C. G. (2005). Commercial mushrooms and bean sprouts are a source of *Pseudomonas aeruginosa*. *J Clin Microbiol* **43**, 5830-5831.

D'Argenio, D. A., Wu, M., Hoffman, L. R. & other authors (2007). Growth phenotypes of *Pseudomonas aeruginosa lasR* mutants adapted to the airways of cystic fibrosis patients. *Mol Microbiol* **64**, 512-533.

Darwin, C. (1859). *On the origin of Species*.

Davidson, S., Yellin, E. O., Shaked, I. & Rubinstein, E. (1982). *Pseudomonas aeruginosa* meningitis treated with an azlocillin combination. *Infection* **10 Suppl 3**, S168-169.

Davies, D. G., Chakrabarty, A. M. & Geesey, G. G. (1993). Exopolysaccharide production in biofilms: substratum activation of alginate gene expression by *Pseudomonas aeruginosa*. *Appl Environ Microbiol* **59**, 1181-1186.

Davies, D. G. & Geesey, G. G. (1995). Regulation of the alginate biosynthesis gene *algC* in *Pseudomonas aeruginosa* during biofilm development in continuous culture. *Appl Environ Microbiol* **61**, 860-867.

Davies, D. G., Parsek, M. R., Pearson, J. P., Iglewski, B. H., Costerton, J. W. & Greenberg, E. P. (1998). The Involvement of Cell-to-Cell Signals in the Development of a Bacterial Biofilm. *Science* **280**, 295-298.

Dayton, S. L., Blasi, D., Chipps, D. D. & Smith, R. F. (1974). Epidemiological tracing of *Pseudomonas aeruginosa*: antibiogram and serotyping. *Appl Microbiol* **27**, 1167-1169.

De Vos, D., Lim, A., Jr, Pirnay, J., Struelens, M., Vandenvelde, C., Duinslaeger, L., Vanderkelen, A. & Cornelis, P. (1997). Direct detection and identification of *Pseudomonas aeruginosa* in clinical samples such as skin biopsy specimens and expectorations by multiplex PCR based on two outer membrane lipoprotein genes, *oprI* and *oprL*. *J Clin Microbiol* **35**, 1295-1299.

Debarbieux, L., Leduc, D., Maura, D., Morello, E., Criscuolo, A., Grossi, O., Balloy, V. & Touqui, L. (2010). Bacteriophages can treat and prevent *Pseudomonas aeruginosa* lung infections. *J Infect Dis* **201**, 1096-1104.

Denamur, E., Picard, B., Decoux, G., Denis, J. B. & Elion, J. (1993). The absence of correlation between allozyme and rrn RFLP analysis indicates a high gene flow rate within human clinical *Pseudomonas aeruginosa* isolates. *FEMS Microbiol Lett* **110**, 275-280.

Dipippo, J. L., Nesbo, C. L., Dahle, H., Doolittle, W. F., Birkland, N. K. & Noll, K. M. (2009). *Kosmotoga olearia* gen. nov., sp. nov., a thermophilic, anaerobic heterotroph isolated from an oil production fluid. *Int J Syst Evol Microbiol* **59**, 2991-3000.

Dobzhansky, T. (1937). *Genetics and the origin of species*. New York: Columbia University Press.

Doolittle, W. F. & Papke, R. T. (2006). Genomics and the bacterial species problem. *Genome Biol* **7**, 116.

Doolittle, W. F. (2008). Microbial evolution: stalking the wild bacterial species. *Curr Biol* **18**, R565-567.

Doolittle, W. F. (2009). Eradicating Typological Thinking in Prokaryotic Systematics and Evolution. *Cold Spring Harb Symp Quant Biol*.

Doolittle, W. F. & Zhaxybayeva, O. (2009). On the origin of prokaryotic species. *Genome Res* **19**, 744-756.

Drenkard, E. & Ausubel, F. M. (2002). *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. **416**, 740-743.

Dubnau, D. (1999). DNA uptake in bacteria. *Annu Rev Microbiol* **53**, 217-244.

Euzeby, J. P. (1997). List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *Int J Syst Bacteriol* **47**, 590-592.

Excoffier, L., Laval, G. & Schneider, S. (2007). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **2005**, 0.

Falush, D., Kraft, C., Taylor, N. S., Correa, P., Fox, J. G., Achtman, M. & Suerbaum, S. (2001). Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. *Proc Natl Acad Sci U S A* **98**, 15056-15061.

Favero, M. S., Carson, L. A., Bond, W. W. & Petersen, N. J. (1971). *Pseudomonas aeruginosa*: growth in distilled water from hospitals. *Science* **173**, 836-838.

Feil, E. J., Smith, J. M., Enright, M. C. & Spratt, B. G. (2000). Estimating recombinational parameters in *Streptococcus pneumoniae* from multilocus sequence typing data. *Genetics* **154**, 1439-1450.

Feil, E. J. & Spratt, B. G. (2001). Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* **55**, 561-590.

Feltman, H., Schulert, G., Khan, S., Jain, M., Peterson, L. & Hauser, A. R. (2001). Prevalence of type III secretion genes in clinical and environmental isolates of *Pseudomonas aeruginosa*. *Microbiology* **147**, 2659-2669.

Filipchenko, Y. (1927). Variabilität und Variation. *Russian eugenical journal* **5**.

Finnan, S., Morrissey, J. P., O'Gara, F. & Boyd, E. F. (2004). Genome diversity of *Pseudomonas aeruginosa* isolates from cystic fibrosis patients and the hospital environment. *J Clin Microbiol* **42**, 5783-5792.

Fisher, M. W., Devlin, H. B. & Gnabasik, F. J. (1969). New immunotype schema for *Pseudomonas aeruginosa* based on protective antigens. *J Bacteriol* **98**, 835-836.

Foght, J. M., Westlake, D. W., Johnson, W. M. & Ridgway, H. F. (1996). Environmental gasoline-utilizing isolates and clinical isolates of *Pseudomonas aeruginosa* are taxonomically indistinguishable by chemotaxonomic and molecular techniques. *Microbiology* **142 (Pt 9)**, 2333-2340.

Forterre, P. (2006). The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res* **117**, 5-16.

Fothergill, J., Panagea, S., Hart, C., Walshaw, M., Pitt, T. & Winstanley, C. (2007). Widespread pyocyanin over-production among isolates of a cystic fibrosis epidemic strain. *BMC Microbiology* **7**, 45.

Fraser, C., Hanage, W. P. & Spratt, B. G. (2005). Neutral microepidemic evolution of bacterial pathogens. *Proc Natl Acad Sci U S A* **102**, 1968-1973.

Fraser, C., Hanage, W. P. & Spratt, B. G. (2007). Recombination and the nature of bacterial speciation. *Science* **315**, 476-480.

Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G. & Hanage, W. P. (2009). The bacterial species challenge: making sense of genetic and ecological diversity. *Science* **323**, 741-746.

Fraser, G. M. & Hughes, C. (1999). Swarming motility. *Current Opinion in Microbiology* **2**, 630-635.

Giammanco, A., Di Stefano, R., Arista, S., Sinatra, A. & Chiarini, A. (1985). Infections caused by *Pseudomonas aeruginosa*: relatively frequent isolation of serogroup 12 from clinical specimens. *Eur J Epidemiol* **1**, 104-109.

Goldberg, J. B. & Pier, G. B. (2000). The role of the CFTR in susceptibility to *Pseudomonas aeruginosa* infections in cystic fibrosis. *Trends Microbiol* **8**, 514-520.

Gooderham, W. J. & Hancock, R. E. (2009). Regulation of virulence and antibiotic resistance by two-component regulatory systems in *Pseudomonas aeruginosa*. *FEMS Microbiol Rev* **33**, 279-294.

Gould, S. J. (2002). *The Structure of Evolutionary Theory*: Harvard: Belknap Harvard.

Govan, J. R. & Nelson, J. W. (1993). Microbiology of cystic fibrosis lung infections: themes and issues. *J R Soc Med* **86 Suppl 20**, 11-18.

Govan, J. R. & Deretic, V. (1996). Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol Rev* **60**, 539-574.

Green, S. K., Schroth, M. N., Cho, J. J., Kominos, S. K. & Vitanza-jack, V. B. (1974). Agricultural plants and soil as a reservoir for *Pseudomonas aeruginosa*. *Appl Microbiol* **28**, 987-991.

Guttman, D. S. & Dykhuizen, D. E. (1994). Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**, 1380-1383.

Habs, I. (1957). Research on O-antigens of *Pseudomonas aeruginosa*. *Z Hyg Infektionskr* **144**, 218-228.

Hagelueken, G., Adams, T. M., Wiehlmann, L., Widow, U., Kolmar, H., Tummler, B., Heinz, D. W. & Schubert, W. D. (2006). The crystal structure of SdsA1, an alkylsulfatase from *Pseudomonas aeruginosa*, defines a third class of sulfatases. *Proc Natl Acad Sci U S A* **103**, 7631-7636.

Hanage, W. P., Fraser, C. & Spratt, B. G. (2006a). Sequences, sequence clusters and bacterial species. *Philos Trans R Soc Lond B Biol Sci* **361**, 1917-1927.

Hanage, W. P., Fraser, C. & Spratt, B. G. (2006b). The impact of homologous recombination on the generation of diversity in bacteria. *J Theor Biol* **239**, 210-219.

Hancock, R. E., Mutharia, L. M., Chan, L., Darveau, R. P., Speert, D. P. & Pier, G. B. (1983). *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis: a class of serum-sensitive, nontypable strains deficient in lipopolysaccharide O side chains. *Infect Immun* **42**, 170-177.

Hartl, D. L. & and Clark, A. G. (2007). *Principles of Population Genetics.*: Sunderland, Massachusetts: Sinauer Associates.

Hasanuzzaman, M., Umadhay-Briones, K. M., Zsiros, S. M., Morita, N., Nodasaka, Y., Yumoto, I. & Okuyama, H. (2004). Isolation, identification, and characterization of a novel, oil-degrading bacterium, *Pseudomonas aeruginosa* T1. *Curr Microbiol* **49**, 108-114.

Hassan, H. M. & Fridovich, I. (1980). Mechanism of the antibiotic action pyocyanine. *J Bacteriol* **141**, 156-163.

Hassett, D. J., Charniga, L., Bean, K., Ohman, D. E. & Cohen, M. S. (1992). Response of *Pseudomonas aeruginosa* to pyocyanin: mechanisms of resistance, antioxidant defenses, and demonstration of a manganese-cofactored superoxide dismutase. *Infect Immun* **60**, 328-336.

Haubold, B. & Hudson, R. R. (2000). LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**, 847-849.

Hector, J. S. & Johnson, A. R. (1990). Determination of genome size of *Pseudomonas aeruginosa* by PFGE: analysis of restriction fragments. *Nucleic Acids Res* **18**, 3171-3174.

Hentzer M, E. L., Nielsen J, Givskov M. (2003). Quorum sensing: a novel target for the treatment of biofilm infections. *BioDrugs* **17**, 241-250.

Herbeck, J. T., Funk, D. J., Degnan, P. H. & Wernegreen, J. J. (2003). A conservative test of genetic drift in the endosymbiotic bacterium *Buchnera*: slightly deleterious mutations in the chaperonin *groEL*. *Genetics* **165**, 1651-1660.

Herberich, E., Sikorski, J. & Hothorn, T. (2010). A Robust Procedure for Comparing Multiple Means under Heteroscedasticity in Unbalanced Designs. *PLoS ONE* **5**, e9788.

Hobden, J. A., Masinick-McClellan, S., Barrett, R. P., Bark, K. S. & Hazlett, L. D. (1999). *Pseudomonas aeruginosa* Keratitis in Knockout Mice Deficient in Intercellular Adhesion Molecule 1. *Infect Immun* **67**, 972-975.

Hogardt, M., Hoboth, C., Schmoldt, S., Henke, C., Bader, L. & Heesemann, J. (2007). Stage-specific adaptation of hypermutable *Pseudomonas aeruginosa* isolates during chronic pulmonary infection in patients with cystic fibrosis. *J Infect Dis* **195**, 70-80.

Homma, J. Y., Shionoya, H., Yamada, H. & Kawabe, Y. (1971). Production of antibody against *Pseudomonas aeruginosa* and its serological typing. *Jpn J Exp Med* **41**, 89-94.

Howe, T. R. & Iglewski, B. H. (1984). Isolation and characterization of alkaline protease-deficient mutants of *Pseudomonas aeruginosa* in vitro and in a mouse eye model. *Infect Immun* **43**, 1058-1063.

Hoyle BD, C. J. (1991). Bacterial resistance to antibiotics: the role of biofilms. *Prog Drug Res* **37**, 91-105.

Hsueh, P. R., Teng, L. J., Yang, P. C., Chen, Y. C., Ho, S. W. & Luh, K. T. (1998). Persistence of a multidrug-resistant *Pseudomonas aeruginosa* clone in an intensive care burn unit. *J Clin Microbiol* **36**, 1347-1351.

Hughes, K. A., Sutherland, I. W. & Jones, M. V. (1998). Biofilm susceptibility to bacteriophage attack: the role of phage-borne polysaccharide depolymerase. *Microbiology* **144 (Pt 11)**, 3039-3047.

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J. & Polz, M. F. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**, 1081-1085.

Huxley, J. S. (1942). *Evolution: The Modern Synthesis*: Allen and Unwin, London.

Jain, M., Ramirez, D., Seshadri, R. & other authors (2004). Type III secretion phenotypes of *Pseudomonas aeruginosa* strains change during infection of individuals with cystic fibrosis. *J Clin Microbiol* **42**, 5229-5237.

Kayser W. V., Hickman K. C. D., Bond W. W., Favero M. S. & A., a. C. L. (1975). Bacteriological evaluation of an ultra-pure water-distilling system. *Appl Microbiol* **30**, 704–706.

Kernacki, K. A., Hobden, J. A., Hazlett, L. D., Fridman, R. & Berk, R. S. (1995). In vivo bacterial protease production during *Pseudomonas aeruginosa* corneal infection. *Invest Ophthalmol Vis Sci* **36**, 1371-1378.

Kessler, E., Safrin, M., Olson, J. C. & Ohman, D. E. (1993). Secreted *LasA* of *Pseudomonas aeruginosa* is a staphylolytic protease. *J Biol Chem* **268**, 7503-7508.

Khan, N. H., Ishii, Y., Kimata-Kino, N., Esaki, H., Nishino, T., Nishimura, M. & Kogure, K. (2007). Isolation of *Pseudomonas aeruginosa* from Open Ocean and Comparison with Freshwater, Clinical, and Animal Isolates. *Microb Ecol*.

Khan, N. H., Ahsan, M., Yoshizawa, S., Hosoya, S., Yokota, A. & Kogure, K. (2008). Multilocus sequence typing and phylogenetic analyses of *Pseudomonas aeruginosa* Isolates from the ocean. *Appl Environ Microbiol* **74**, 6194-6205.

Kiewitz, C. & Tummler, B. (2000). Sequence diversity of *Pseudomonas aeruginosa*: impact on population structure and genome evolution. *J Bacteriol* **182**, 3125-3135.

Kimata, N., Nishino, T., Suzuki, S. & Kogure, K. (2004). *Pseudomonas aeruginosa* isolated from marine environments in Tokyo Bay. *Microb Ecol* **47**, 41-47.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature* **217**, 624–626.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.*

Knezevic, P. & Petrovic, O. (2008). A colorimetric microtiter plate method for assessment of phage effect on *Pseudomonas aeruginosa* biofilm. *J Microbiol Methods* **74**, 114-118.

Knezevic, P., Kostanjsek, R., Obreht, D. & Petrovic, O. (2009). Isolation of *Pseudomonas aeruginosa* Specific Phages with Broad Activity Spectra. *Curr Microbiol* **Volume 59**, 173-180.

Koeppel, A., Perry, E. B., Sikorski, J. & other authors (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci U S A* **105**, 2504-2509.

Konstantinidis, K. T. & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* **102**, 2567-2572.

Koonin, E. V. (2009). Darwinian evolution in the light of genomics. *Nucleic Acids Res* **37**, 1011-1034.

Kuo, C. H., Moran, N. A. & Ochman, H. (2009). The consequences of genetic drift for bacterial genome complexity. *Genome Res* **19**, 1450-1454.

Lanyi, B. (1966). Serological properties of *Pseudomonas aeruginosa*. I. Group-specific somatic antigens. *Acta Microbiol Acad Sci Hung* **13**, 295-318.

Lapage, S. P., Sneath, P. H. A., Lessel, E. F., Skerman, V. B. D., Seeliger, H. P. R. & Clark, W. A. (1992). *International Code of Nomenclature of Bacteria (1990 Revision). Bacteriological Code*: Washington, DC: American Society for Microbiology.

Lee, B., Haagensen, J. A., Ciofu, O., Andersen, J. B., Hoiby, N. & Molin, S. (2005). Heterogeneity of biofilms formed by nonmucoid *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. *J Clin Microbiol* **43**, 5247-5255.

Lee, D. G., Urbach, J. M., Wu, G. & other authors (2006). Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol* **7**, R90.

Legnani, P., Leoni, E., Rapuano, S., Turin, D. & Valenti, C. (1999). Survival and growth of *Pseudomonas aeruginosa* in natural mineral water: a 5-year study. *Int J Food Microbiol* **53**, 153-158.

Leone I., C. M. G., Raso T., Zucca M. and Savoia D. (2008). Phenotypic and genotypic characterization of *Pseudomonas aeruginosa* from cystic fibrosis patients. *European Journal of Clinical Microbiology & Infectious Diseases* **27**, 1093-1099.

Lewis, D. A., Jones, A., Parkhill, J., Speert, D. P., Govan, J. R., Lipuma, J. J., Lory, S., Webb, A. K. & Mahenthiralingam, E. (2005). Identification of DNA markers for a transmissible Pseudomonas aeruginosa cystic fibrosis strain. *Am J Respir Cell Mol Biol* **33**, 56-64.

Lewis, K. (2001). Riddle of biofilm resistance. *Antimicrob Agents Chemother* **45**, 999-1007.

Lin, H.-H., Huang, S.-p., Teng, H.-C., Ji, D.-D., Chen, Y.-S. & Chen, Y.-L. (2006). Presence of the *exoU* Gene of *Pseudomonas aeruginosa* Is Correlated with Cytotoxicity in MDCK Cells but Not with Colonization in BALB/c Mice. *J Clin Microbiol* **44**, 4596-4597.

Liu, P. V., Matsumoto, H., Kusama, H. & Bergan, T. (1983). Survey of Heat-Stable, Major Somatic Antigens of *Pseudomonas aeruginosa*. *Int J Syst Bacteriol* **33**, 256-264.

Luzar, M. A. & Montie, T. C. (1985). Avirulence and altered physiological properties of cystic fibrosis strains of *Pseudomonas aeruginosa*. *Infect Immun* **50**, 572-576.

Luzar, M. A., Thomassen, M. J. & Montie, T. C. (1985). Flagella and motility alterations in *Pseudomonas aeruginosa* strains from patients with cystic fibrosis: relationship to patient clinical condition. *Infect Immun* **50**, 577-582.

Lynch, M. & Conery, J. S. (2003). The origins of genome complexity. *Science* **302**, 1401-1404.

Lynch, M. (2006). Streamlining and simplification of microbial genome architecture. *Annu Rev Microbiol* **60**, 327-349.

Lynch, M. (2007a). The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences* **104**, 8597-8604.

Lynch, M. (2007b). *The Origins of Genome Architecture*.

Mahenthiralingam, E., Campbell, M. E. & Speert, D. P. (1994). Nonmotility and phagocytic resistance of *Pseudomonas aeruginosa* isolates from chronically colonized patients with cystic fibrosis. *Infect Immun* **62**, 596-605.

Mahenthiralingam, E. & Speert, D. (1995). Nonopsonic phagocytosis of *Pseudomonas aeruginosa* by macrophages and polymorphonuclear leukocytes requires the presence of the bacterial flagellum. *Infect Immun* **63**, 4519-4523.

Maiden, M. C., Bygraves, J. A., Feil, E. & other authors (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* **95**, 3140-3145.

Maiden, M. C. (2006). Multilocus sequence typing of bacteria. *Annu Rev Microbiol* **60**, 561-588.

Majewski, J. (2001). Sexual isolation in bacteria. *FEMS Microbiol Lett* **199**, 161-169.

Makin, S. A. & Beveridge, T. J. (1996). *Pseudomonas aeruginosa* PAO1 ceases to express serotype-specific lipopolysaccharide at 45 degrees C. *J Bacteriol* **178**, 3350-3352.

Mallet, M. (1995). A species definition for the modern synthesis. *Trends Ecol Evol* **10**, 294-299.

Margos, G., Gatewood, A. G., Aanensen, D. M. & other authors (2008). MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. *Proceedings of the National Academy of Sciences* **105**, 8730-8735.

Matic, I., Taddei, F. & Radman, M. (1996). Genetic barriers among bacteria. *Trends Microbiol* **4**, 69-72.

Mayr, E. (1944). *Systematics and the Origin of Species.* New York: Columbia University Press.

Mayr, E. (1997). The objects of selection. *Proc Natl Acad Sci U S A* **94**, 2091-2094.

McVay, C. S., Velasquez, M. & Fralick, J. A. (2007). Phage therapy of *Pseudomonas aeruginosa* infection in a mouse burn wound model. *Antimicrob Agents Chemother* **51**, 1934-1938.

Meitert, T. (1964). Contribution to the study of the antigenic structure of pyocyanic bacilli (*Pseudomonas aeruginosa*). II. Separation of serological groups by means of the "O" antigen. *Arch Roum Pathol Exp Microbiol* **23**, 679-688.

Mena, K. D. & Gerba, C. P. (2009). Risk assessment of *Pseudomonas aeruginosa* in water. *Rev Environ Contam Toxicol* **201**, 71-115.

Merabishvili, M., Pirnay, J. P., Verbeken, G. & other authors (2009). Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *PLoS One* **4**, e4944.

Merod, R. T., Warren, J. E., McCaslin, H. & Wuertz, S. (2007). Toward Automated Analysis of Biofilm Architecture: Bias Caused by Extraneous Confocal Laser Scanning Microscopy Images. *Appl Environ Microbiol* **73**, 4922-4930.

Morales, G., Wiehlmann, L., Gudowius, P., van Delden, C., Tummler, B., Martinez, J. L. & Rojo, F. (2004). Structure of *Pseudomonas aeruginosa* populations analyzed by single nucleotide polymorphism and pulsed-field gel electrophoresis genotyping. *J Bacteriol* **186**, 4228-4237.

Moran, N. & Mira, A. (2001). The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. *Genome Biology* **2**, 0054.0051 - 0054.0012.

Moran, X., Estrada, M., Gasol, J. & Pedros-Alio, C. (2002). Dissolved primary production and the strength of phytoplankton bacterioplankton coupling in contrasting marine regions. *Microb Ecol* **44**, 217 - 223.

Mueller, L., de Brouwer, J., Almeida, J., Stal, L. & Xavier, J. (2006). Analysis of a marine phototrophic biofilm by confocal laser scanning microscopy using the new image quantification software PHLIP. *BMC Ecology* **6**, 1.

Mutch, L. A. & Young, J. P. (2004). Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes. *Mol Ecol* **13**, 2435-2444.

Mutharia, L. M., Nicas, T. I. & Hancock, R. E. (1982). Outer membrane proteins of *Pseudomonas aeruginosa* serotype strains. *J Infect Dis* **146**, 770-779.

Nicas, T. I. & Iglewski, B. H. (1985). The contribution of exoproducts to virulence of *Pseudomonas aeruginosa*. *Can J Microbiol* **31**, 387-392.

Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299-304.

O'Flaherty, S., Ross, R. P. & Coffey, A. (2009). Bacteriophage and their lysins for elimination of infectious bacteria. *FEMS Microbiology Reviews* **33**, 801-819.

Ohman, D. E., Cryz, S. J. & Iglewski, B. H. (1980). Isolation and characterization of *Pseudomonas aeruginosa* PAO mutant that produces altered elastase. *J Bacteriol* **142**, 836-842.

Ohta, T. (1973). Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96-98.

Olsen, G. J., Woese, C. R. & Overbeek, R. (1994). The winds of (evolutionary) change: breathing new life into microbiology. *J Bacteriol* **176**, 1-6.

Orr, H. A. (2005). The genetic theory of adaptation: a brief history. **6**, 119-127.

O'Toole, G., Kaplan, H. B. & Kolter, R. (2000). Biofilm formation as microbial development. *Annual Review of Microbiology* **54**, 49-79.

O'Toole, G. A. & Kolter, R. (1998a). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol Microbiol* **30**, 295-304.

O'Toole, G. A. & Kolter, R. (1998b). Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Molecular Microbiology* **28**, 449-461.

Ozaki, K., Inoue, A., Atobe, H., Takahashi, E. & Konishi, S. (1990). Serotypes and antimicrobial susceptibility of *Pseudomonas aeruginosa* strains isolated from diseased dogs. *Nippon Juigaku Zasshi* **52**, 233-239.

Palmer, K. L., Mashburn, L. M., Singh, P. K. & Whiteley, M. (2005). Cystic fibrosis sputum supports growth and cues key aspects of *Pseudomonas aeruginosa* physiology. *J Bacteriol* **187**, 5267-5277.

Papke, R. T., Ramsing, N. B., Bateson, M. M. & Ward, D. M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5**, 650-659.

Papke, R. T. & Ward, D. M. (2004). The importance of physical isolation to microbial diversification. *FEMS Microbiol Ecol* **48**, 293-303.

Patzer, J. & Dzierzanowska, D. (1994). The incidence of serotype O12 and multiresistance amongst *Pseudomonas aeruginosa* clinical isolates. *J Antimicrob Chemother* **34**, 165-170.

Pavlovskis, O. R. & Wretlind, B. (1979). Assessment of protease (elastase) as a *Pseudomonas aeruginosa* virulence factor in experimental mouse burn infection. *Infect Immun* **24**, 181-187.

Pedersen, S. S., Jensen, T., Pressler, T., Hoiby, N. & Rosendal, K. (1986a). Does centralized treatment of cystic fibrosis increase the risk of *Pseudomonas aeruginosa* infection? *Acta Paediatr Scand* **75**, 840-845.

Pedersen, S. S., Koch, C., Hoiby, N. & Rosendal, K. (1986b). An epidemic spread of multiresistant *Pseudomonas aeruginosa* in a cystic fibrosis centre. *J Antimicrob Chemother* **17**, 505-516.

Pellett, S., Bigley, D. V. & Grimes, D. J. (1983). Distribution of *Pseudomonas aeruginosa* in a riverine ecosystem. *Appl Environ Microbiol* **45**, 328-332.

Pham, T. H., Webb, J. S. & Rehm, B. H. A. (2004). The role of polyhydroxyalkanoate biosynthesis by *Pseudomonas aeruginosa* in rhamnolipid and alginate production as well as stress tolerance and biofilm formation. *Microbiology* **150**, 3405-3413.

Picard, B., Denamur, E., Barakat, A., Elion, J. & Goullet, P. (1994). Genetic heterogeneity of *Pseudomonas aeruginosa* clinical isolates revealed by esterase electrophoretic polymorphism and restriction fragment length polymorphism of the ribosomal RNA gene region. *J Med Microbiol* **40**, 313-322.

Pirnay, J. P., De Vos, D., Cochez, C., Bilocq, F., Vanderkelen, A., Zizi, M., Ghysels, B. & Cornelis, P. (2002). *Pseudomonas aeruginosa* displays an epidemic population structure. *Environ Microbiol* **4**, 898-911.

Pirnay, J. P., Matthijs, S., Colak, H. & other authors (2005). Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. *Environ Microbiol* **7**, 969-980.

Pirnay, J.-P., Bilocq, F., Pot, B. & other authors (2009). *Pseudomonas aeruginosa* Population Structure Revisited. *PLoS ONE* **4**, e7740.

Pitt, T. L., MacDougall, J., Penketh, A. R. & Cooke, E. M. (1986). Polyagglutinating and non-typable strains of *Pseudomonas aeruginosa* in cystic fibrosis. *J Med Microbiol* **21**, 179-186.

Polz, M. F., Hunt, D. E., Preheim, S. P. & Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos Trans R Soc Lond B Biol Sci* **361**, 2009-2021.

Poole, K. (2001). Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *J Mol Microbiol Biotechnol* **3**, 255-264.

Pratt, L. A. & Kolter, R. (1998). Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Molecular Microbiology* **30**, 285-293.

Rakhimova, E., Wiehlmann, L., Brauer, A. L., Sethi, S., Murphy, T. F. & Tummler, B. (2009). *Pseudomonas aeruginosa* population biology in chronic obstructive pulmonary disease. *The Journal of Infectious Diseases* **200**, 1928-1935.

Rashid, M. H. & Kornberg, A. (2000). Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **97**, 4885-4890.

Rasko, D. A., Rosovitz, M. J., Okstad, O. A., Fouts, D. E., Jiang, L., Cer, R. Z., Kolsto, A. B., Gill, S. R. & Ravel, J. (2007). Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the *B. cereus*-group plasmids, including *Bacillus anthracis* pXO1. *J Bacteriol* **189**, 52-64.

Reyes, E. A., Bale, M. J., Cannon, W. H. & Matsen, J. M. (1981). Identification of *Pseudomonas aeruginosa* by pyocyanin production on Tech agar. *J Clin Microbiol* **13**, 456-458.

Richter, M. & Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences* **106**, 19126-19131.

Romling, U., Wingender, J., Muller, H. & Tummler, B. (1994). A major *Pseudomonas aeruginosa* clone common to patients and aquatic habitats. *Appl Environ Microbiol* **60**, 1734-1738.

Romling, U., Greipel, J. & Tummler, B. (1995). Gradient of genomic diversity in the *Pseudomonas aeruginosa* chromosome. *Mol Microbiol* **17**, 323-332.

Romling, U., Kader, A., Sriramulu, D. D., Simm, R. & Kronvall, G. (2005). Worldwide distribution of *Pseudomonas aeruginosa* clone C strains in the aquatic environment and cystic fibrosis patients. *Environ Microbiol* **7**, 1029-1038.

Rose, M. R. & Oakley, T. H. (2007). The new biology: beyond the Modern Synthesis. *Biol Direct* **2**.

Rossello-Mora, R. & Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews* **25**.

Ruimy, R., Genauzeau, E., Barnabe, C., Beaulieu, A., Tibayrenc, M. & Andremont, A. (2001). Genetic diversity of *Pseudomonas aeruginosa* strains isolated from ventilated patients with nosocomial pneumonia, cancer patients with bacteremia, and environmental water. *Infect Immun* **69**, 584-588.

Sadikot, R. T., Blackwell, T. S., Christman, J. W. & Prince, A. S. (2005). Pathogen-host interactions in *Pseudomonas aeruginosa* pneumonia. *Am J Respir Crit Care Med* **171**, 1209-1223.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning: a Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*.

Sander, R. (2001). Otitis externa: a practical guide to treatment and prevention. *Am Fam Physician* **63**, 927-936, 941-922.

Sandvik, O. (1960). Serological comparison between strains of *Pseudomonas aeruginosa* from human and animal sources. *Acta Pathol Microbiol Scand* **48**, 56-60.

Schoental, R. (1941). The nature of the antibacterial agents present in *Pseudomonas pyocyanea* cultures. *BritJ Exp Pathol* **22**, 137.

Schroeter, J. (1872). Über einige durch Bacterien gebildete Pigmente. *Beitr Biol Pflanz*, 109-126.

Schulert, G. S., Feltman, H., Rabin, S. D., Martin, C. G., Battle, S. E., Rello, J. & Hauser, A. R. (2003). Secretion of the toxin *ExoU* is a marker for highly virulent *Pseudomonas aeruginosa* isolates obtained from patients with hospital-acquired pneumonia. *J Infect Dis* **188**, 1695-1706.

Schuster, M., Lostroh, C. P., Ogi, T. & Greenberg, E. P. (2003). Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J Bacteriol* **185**, 2066-2079.

Scott, F. & Pitt, T. (2004). Identification and characterization of transmissible Pseudomonas aeruginosa strains in cystic fibrosis patients in England and Wales. *J Med Microbiol* **53**, 609 - 615.

Sener, B., Koseoglu, O., Ozcelik, U., Kocagoz, T. & Gunalp, A. (2001). Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic fibrosis. *Int J Med Microbiol* **291**, 387-393.

Shaver, C. M. & Hauser, A. R. (2004). Relative contributions of *Pseudomonas aeruginosa* ExoU, ExoS, and ExoT to virulence in the lung. *Infect Immun* **72**, 6969-6977.

Shigemura, K., Arakawa, S., Sakai, Y., Kinoshita, S., Tanaka, K. & Fujisawa, M. (2006). Complicated urinary tract infection caused by *Pseudomonas aeruginosa* in a single institution. *International Journal of Urology* **13**, 538-542.

Sikorski, J. & Nevo, E. (2005). Adaptation and incipient sympatric speciation of Bacillus simplex under microclimatic contrast at "Evolution Canyons" I and II, Israel. *Proc Natl Acad Sci U S A* **102**, 15924-15929.

Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York: Columbia University Press.

Sjogren, R. E. & Gibson, M. J. (1981). Bacterial survival in a dilute environment. *Appl Environ Microbiol* **41**, 1331-1336.

Skerman, V. B. D., McGowan, V. & Sneath, P. H. A. (1980). Approved Lists of Bacterial Names. *Int J Syst Bacteriol* **30**, 225-420.

Smith, E. E., Buckley, D. G., Wu, Z. & other authors (2006a). Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A* **103**, 8487-8492.

Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., Gordon, S. V., van Soolingen, D., Hewinson, R. G. & Smith, J. M. (2006b). Ecotypes of the *Mycobacterium tuberculosis* complex. *J Theor Biol* **239**, 220-225.

Soberón-Chávez, G., Lépine, F. & Déziel, E. (2005). Production of rhamnolipids by *Pseudomonas aeruginosa*. *Appl Microbiol Biotechnol* **68**, 718–725.

Spratt, B. G. & Maiden, M. C. (1999). Bacterial population genetics, evolution and epidemiology. *Philos Trans R Soc Lond B Biol Sci* **354**, 701-710.

Sriramulu, D. D., Nimtz, M. & Romling, U. (2005). Proteome analysis reveals adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis lung environment. *Proteomics* **5**, 3712-3721.

Staley, J. T. & Krieg, N. J. (1984). Classification of prokaryotic organisms: an overview. *In N R Krieg and J G Holt (ed), Bergey's manual of systematic bacteriology, vol 1 The Williams & Wilkins Co, Baltimore*, 1–3.

Staley, J. T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. *Philos Trans R Soc Lond B Biol Sci* **361**, 1899-1909.

Stoodley, P., Sauer, K., Davies, D. G. & Costerton, J. W. (2002). Biofilms as complex differentiated communities. *Annual Review of Microbiology* **56**, 187-209.

Stover, C., Pham, X., Erwin, A. & other authors (2000a). Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. *Nature* **406**, 959 - 964.

Stover, C. K., Pham, X. Q., Erwin, A. L. & other authors (2000b). Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* **406**, 959-964.

Sulakvelidze, A. & Kutter, E. (2005). Bacteriophages Biology and Applications (Kutter E & Sulakvelidze A, eds). *CRC Press, Boca Raton, FL*.

Suter, S. (1994). The role of bacterial proteases in the pathogenesis of cystic fibrosis. *Am J Respir Crit Care Med* **150**, S118-122.

Suthereland, I. W., Hughes, K. A., Skillman, L. C. & Tait, K. (2004). The interaction of phage and biofilms. *FEMS Microbiol Lett* **232**, 1-6.

Suttle, C. A. (2007). Marine viruses--major players in the global ecosystem. *Nat Rev Microbiol* **5**, 801-812.

Szoboszlay, S., Atzel, B. & Kriszt, B. (2003). Comparative biodegradation examination of *Pseudomonas aeruginosa* (ATCC 27853) and other oil degraders on hydrocarbon contaminated soil. *Commun Agric Appl Biol Sci* **68**, 207-210.

Talarmin, A., Dubrous, P., Gerome, P. & Buisson, Y. (1996). Study of *Pseudomonas aeruginosa* serotype O12 isolates with a common antibiotic susceptibility pattern. *Eur J Clin Microbiol Infect Dis* **15**, 459-464.

Thiel, K. (2004). Old dogma, new tricks--21st Century phage therapy. *Nat Biotechnol* **22**, 31-36.

Thomas, C. M. & Nielsen, K. M. (2005). Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nat Rev Micro* **3**, 711-721.

Thompson, J. R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D. E., Benoit, J., Sarma-Rupavtarm, R., Distel, D. L. & Polz, M. F. (2005). Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**, 1311-1313.

Tindall, B. J., Rossello-Mora, R., Busse, H.-J., Ludwig, W. & Kampfer, P. (2009). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* **60**, 249-266.

Tress, B. M., Hellyar, A. G., Pennington, J., Thomson, K. R., Desmond, P. M., Martinkus, J. & Lavan, J. J. (1994). Multiple doses of contrast medium from a single container: bacteriological studies. *Australas Radiol* **38**, 115-118.

Twining, S. S., Kirschner, S. E., Mahnke, L. A. & Frank, D. W. (1993). Effect of *Pseudomonas aeruginosa* elastase, alkaline protease, and exotoxin A on corneal proteinases and proteins. *Invest Ophthalmol Vis Sci* **34**, 2699-2712.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. & J., S. (1996). Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics. *MICROBIOLOGICAL REVIEWS* **60**, 407–438.

Verder, E. & Evans, J. (1961). A proposed antigenic schema for the identification of strains of *Pseudomonas aeruginosa*. *J Infect Dis* **109**, 183-193.

Vives-Flórez, M. & Garnica, D. (2006). Comparison of virulence between clinical and environmental *Pseudomonas aeruginosa* isolates. *Int Microbiol* **9**, 247-252.

Vos, M. & Didelot, X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *Isme J* **3**, 199-208.

Waite, R. D., Papakonstantinopoulou, A., Littler, E. & Curtis, M. A. (2005). Transcriptome analysis of *Pseudomonas aeruginosa* growth: comparison of gene expression in planktonic cultures and developing and mature biofilms. *J Bacteriol* **187**, 6571-6576.

Waite, R. D., Paccanaro, A., Papakonstantinopoulou, A., Hurst, J. M., Saqi, M., Littler, E. & Curtis, M. A. (2006). Clustering of *Pseudomonas aeruginosa* transcriptomes from planktonic cultures, developing and mature biofilms reveals distinct expression profiles. *BMC Genomics* **7**, 162.

Ward, D. M. (1998). A natural species concept for prokaryotes. *Curr Opin Microbiol* **1**, 271-277.

Ward, D. M., Ferris, M. J., Nold, S. C. & Bateson, M. M. (1998). A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol Mol Biol Rev* **62**, 1353-1370.

Ward, D. M., Bateson, M. M., Ferris, M. J., Kuhl, M., Wieland, A., Koeppel, A. & Cohan, F. M. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos Trans R Soc Lond B Biol Sci* **361**, 1997-2008.

Ward, D. M., Cohan, F. M., Bhaya, D., Heidelberg, J. F., Kuhl, M. & Grossman, A. (2008). Genomics, environmental genomics and the issue of microbial species. *Heredity* **100**, 207-219.

Watanabe, R., Matsumoto, T., Sano, G. & other authors (2007). Efficacy of bacteriophage therapy against gut-derived sepsis caused by *Pseudomonas aeruginosa* in mice. *Antimicrob Agents Chemother* **51**, 446-452.

Watanakunakorn, C. (1975). Vertebral osteomyelitis as a complication of *Pseudomonas aeruginosa* pneumonia. *South Med J* **68**, 173-176.

Watine, J. (1999). Is the European interhospital clonal spread of serotype O12 *Pseudomonas aeruginosa* related to the patients' prolonged carriage duration? *Infect Control Hosp Epidemiol* **20**, 460-461.

Wayne, L. G., Brenner, D. J., Colwell, R. R. & other authors (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Int J Syst Bacteriol* **37**, 463 - 464.

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol Rev* **28**, 127-181.

Whitaker, R. J., Grogan, D. W. & Taylor, J. W. (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**, 976-978.

Whitaker, R. J. (2006). Allopatric origins of microbial species. *Philos Trans R Soc Lond B Biol Sci* **361**, 1975-1984.

Whitman, W. B., Coleman, D. C. & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* **95**, 6578-6583.

Wiehlmann, L., Wagner, G., Cramer, N. & other authors (2007). Population structure of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **104**, 8101-8106.

Wilson, R., Sykes, D. A., Watson, D., Rutman, A., Taylor, G. W. & Cole, P. J. (1988). Measurement of *Pseudomonas aeruginosa* phenazine pigments in sputum and assessment of their contribution to sputum sol toxicity for respiratory epithelium. *Infect Immun* **56**, 2515-2517.

Wimpenny, J. W. T. & Colasanti, R. (1997). A unifying hypothesis for the structure of microbial biofilms based on cellular automaton models. *FEMS Microbiology Ecology* **22**, 1-16.

Wolfgang, M. C., Kulasekara, B. R., Liang, X., Boyd, D., Wu, K., Yang, Q., Miyada, C. G. & Lory, S. (2003). Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **100**, 8484-8489.

Yokota, S., Hayashi, T. & Matsumoto, H. (1994). Identification of the lipopolysaccharide core region as the receptor site for a cytotoxin-converting phage, phi CTX, of *Pseudomonas aeruginosa*. *J Bacteriol* **176**, 5262-5269.

7. Appendix

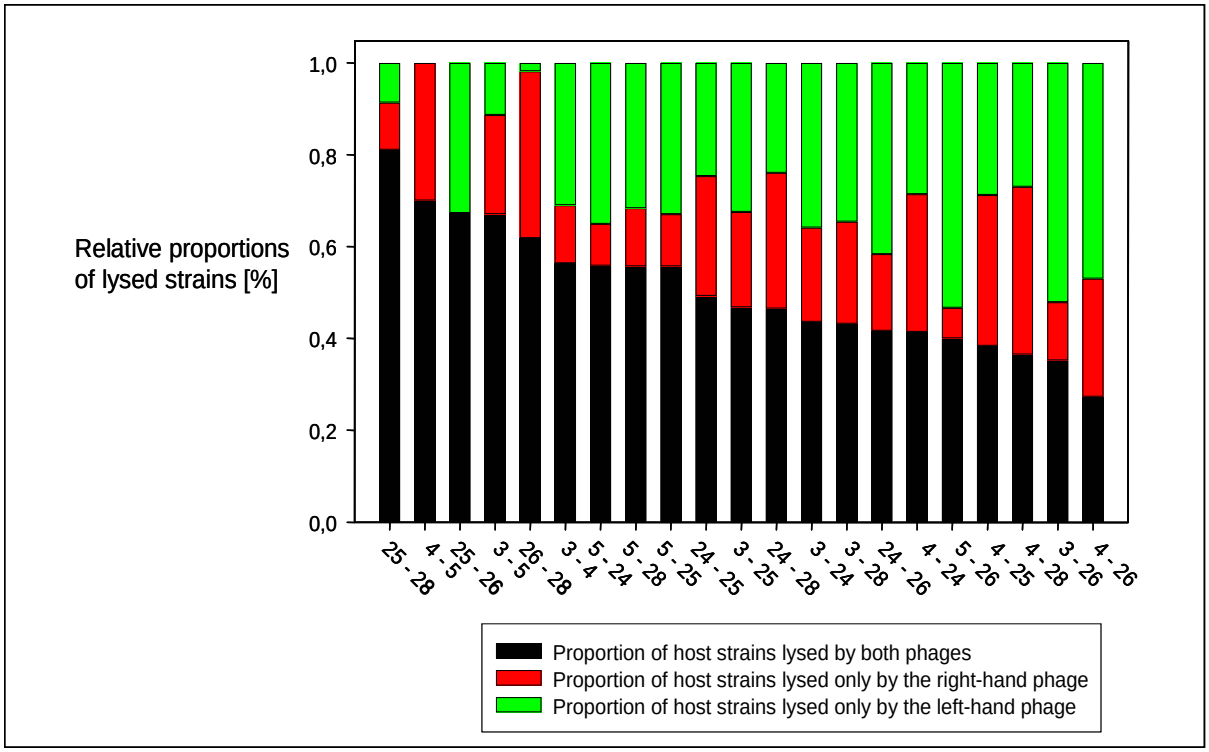

Figure 6a. Pairwise comparison of joint host specificity of phages. On the x-axes names of the paired phages are given: 3 (JG003), 4 (JG004), 5 (JG005), 24 (JG024), 25 (JG025), 26 (JG026) and 28 (JG028).

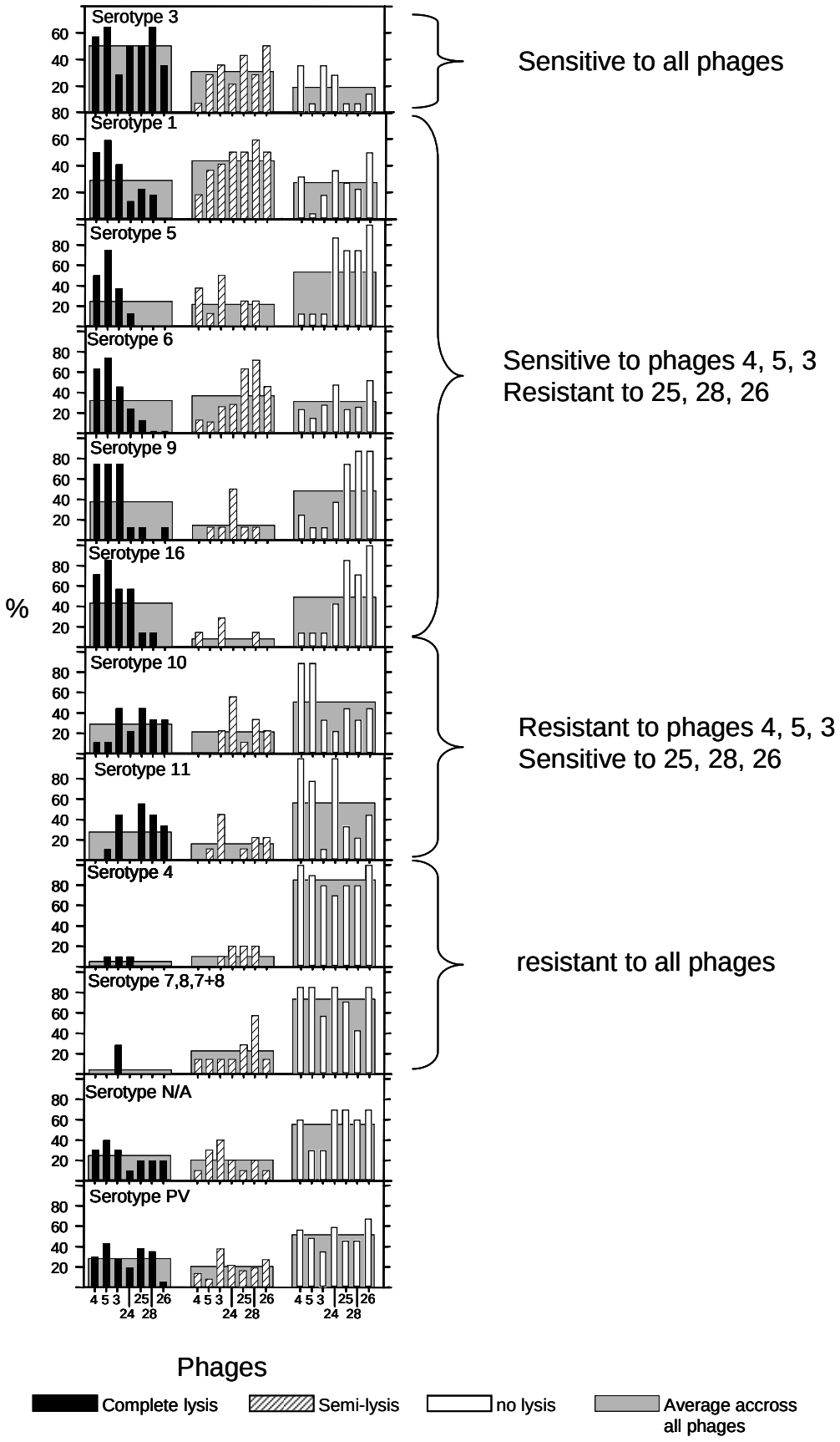

Figure 9. Distribution of lysis efficiencies of the 7 phages across different serotypes. The proportions (%) of the lysis phenotypes for individual phages among all strains with a given serotype are shown as black (clear lysis), striped (semi lysis) or white (no lysis) bars. The results from serotype 7 and 8 are summed up, as these serotypes are frequently not distinguishable. The grey field in the background indicates the average value from the combined data of all phages. The data are based on the combined set of 99 environmental and 96 clinical strains.

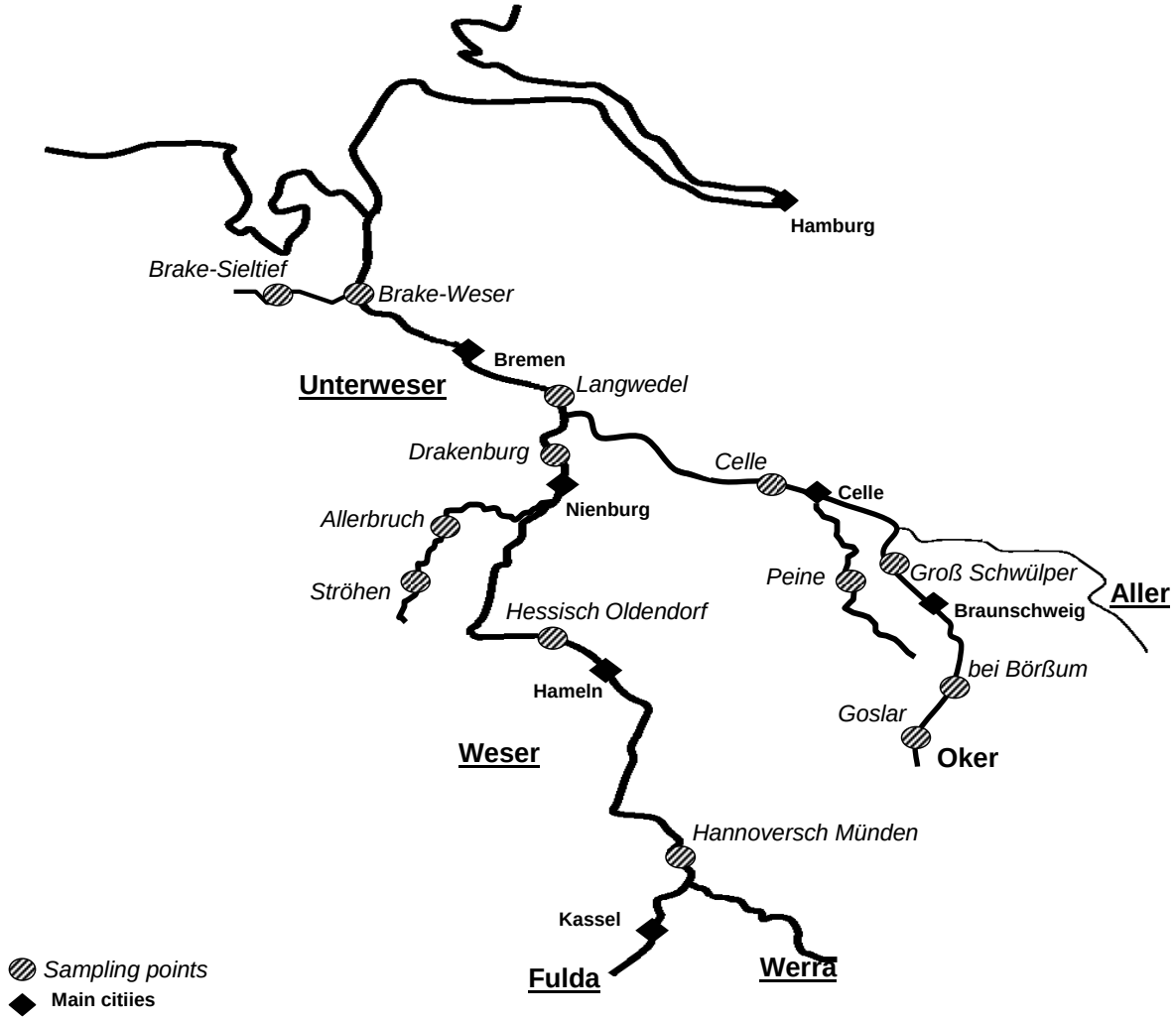

Figure 10. Geographical localization of the 13 sampling places along Lower Saxony rivers.

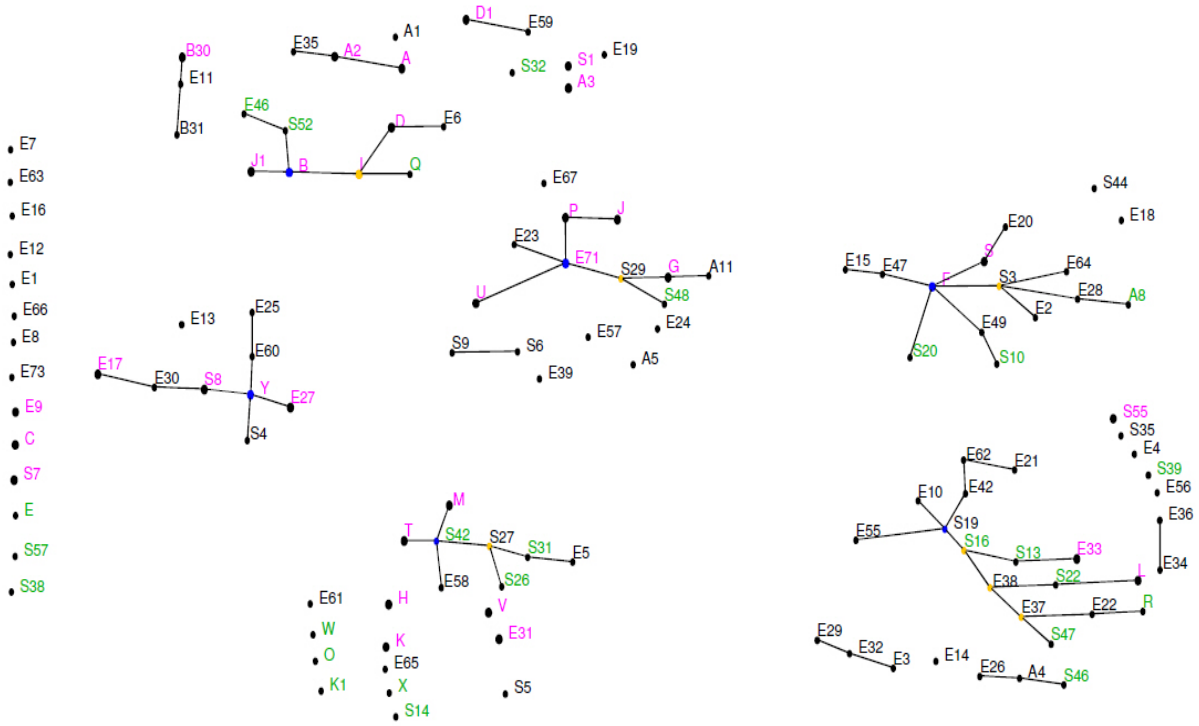

Figure 11. The structure of the genetic relationships among the *P. aeruginosa* host strains as based on eBURST analysis of the core genome SNP pattern. This figure is effectively the same eBURST graph as shown in Figure 1, with the difference that in this figure the names of all clones are given. Pink clone names are represented by both clinical and environmental strains. Uniquely clinical or environmental clones are coloured in green or in black, respectively. The founder strain of the clonal complexes is shown as a blue spot.

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); “-w” denotes winter, “-s” denotes summer.

Ecc group B

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sieltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sieltief_s	**Total number**
A	0	2	0	0	0	1	0	0	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	**7**
A1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
A2	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
A3	0	1	2	1	0	2	0	0	3	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**16**
B	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	**3**
B30	0	0	1	1	0	0	0	0	5	2	0	0	0	5	8	1	0	2	0	1	1	0	0	11	6	**44**
B31	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**3**
D	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	**2**
D1	0	0	1	0	0	0	8	0	0	0	0	1	0	1	0	0	1	2	2	1	0	1	0	0	1	**19**
E11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	**2**
E19	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	**1**
E59	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	**2**
I	0	1	0	0	0	0	0	0	0	0	0	1	1	2	1	0	0	0	0	0	9	0	0	0	0	**15**
J1	0	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**6**
S1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**1**
Total	**0**	**5**	**6**	**3**	**1**	**4**	**11**	**0**	**12**	**9**	**4**	**4**	**2**	**10**	**11**	**1**	**2**	**7**	**3**	**2**	**11**	**1**	**0**	**11**	**7**	**127**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted) also “-w” denotes winter, “-s” denotes summer.

Ecc group F

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sieltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sieltief_s	**Total number**
E15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	**3**
E18	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E2	0	0	0	0	0	0	0	0	0	8	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	**12**
E20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	0	0	0	**6**
E28	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	**1**
E49	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	**2**
E64	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**18**
F	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**1**
S	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
S3	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**
S44	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
Total	**0**	**1**	**3**	**0**	**0**	**3**	**1**	**18**	**0**	**8**	**0**	**0**	**1**	**4**	**5**	**0**	**0**	**2**	**0**	**0**	**1**	**0**	**1**	**0**	**1**	**49**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); "-w" denotes winter, "-s" denotes summer.

Ecc group E71

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sielltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sielltief_s	Total number
A11	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
A5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E23	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
E39	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3
E57	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	3
E67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
E71	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
G	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	9
J	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	4
P	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
S29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
S6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
S9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Total	**0**	**11**	**0**	**1**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**5**	**2**	**2**	**1**	**1**	**1**	**0**	**0**	**2**	**1**	**0**	**2**	**0**	**30**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); "-w" denotes winter, "-s" denotes summer.

Ecc group S19

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sielltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sielltief_s	Total number
A4	0	1	0	0	0	0	0	0	2	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5
E10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
E14	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2
E21	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	3
E22	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	3
E26	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E29	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
E32	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
E33	0	4	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	7
E34	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
E36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
E37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
E38	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
E55	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E56	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	3
E62	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	4
L	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	4
S19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	6
S35	0	0	0	0	0	1	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	4
S55	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Total	**0**	**6**	**1**	**1**	**0**	**2**	**1**	**1**	**6**	**3**	**3**	**1**	**11**	**3**	**3**	**0**	**2**	**1**	**2**	**0**	**2**	**0**	**1**	**0**	**6**	**56**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); “-w” denotes winter, “-s” denotes summer.

Ecc group S42

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sieltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sieltief_s	**Total number**
E31	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**1**
E58	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E61	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	**2**
H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	**1**
K	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	**1**
M	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
S27	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
S5	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**3**
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**1**
V	0	0	0	0	0	0	0	0	2	1	0	13	0	0	1	0	0	0	0	0	0	0	13	0	3	**33**
Total	**0**	**0**	**0**	**1**	**0**	**0**	**1**	**1**	**3**	**3**	**0**	**14**	**0**	**1**	**1**	**0**	**1**	**1**	**0**	**0**	**1**	**0**	**13**	**3**	**3**	**47**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); “-w” denotes winter, “-s” denotes summer.

Ecc group Y

Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sieltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sieltief_s	**Total number**
E13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**1**
E17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**1**
E25	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E27	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**
E30	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
E60	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	**3**
S4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
S8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	**3**
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	**1**
Total	**0**	**1**	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**2**	**0**	**0**	**1**	**2**	**0**	**1**	**1**	**15**

Different colors represent different water qualities (greenish = low polluted, yellowish = moderately polluted, pinkish = strongly polluted); "-w" denotes winter, "-s" denotes summer.

No DLV to the clonal complexes	Haplotype:	Goslar-w	Hann.Muenden	Langwedel-w	Celle-w	Goslar-s	Hann.-Muenden-s	Langwedel-s	Celle-s	Hessisch_Oldendorf_w	Gross_Schwuelper_w	bei_Boerssum_w	Hessisch-Oldendorf-s	Gross-Schwuelper-s	bei-Boerssum-s	Peine-w	Allerbruch-w	Drakenburg-w	Stroehen-w	Brake-Weser-w	Brake-Sieltief-w	Peine-s	Drakenburg-s	Stroehen-s	Brake-Weser_s	Brake-Sieltief_s	Total number
	E7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	E16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	E63	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
	E1	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	19
	E12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
	E66	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3
	E8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	4
	E9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	4
	E73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	C	4	3	1	0	9	2	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	21
	Total	**4**	**3**	**1**	**0**	**9**	**2**	**0**	**0**	**0**	**18**	**0**	**1**	**1**	**0**	**4**	**0**	**3**	**1**	**2**	**0**	**2**	**0**	**0**	**3**	**2**	**56**

Figure 12. Detailed list of the haplotypes' sampling abundancies in every each sampling station, both seasons.

Supplemental Table 1. Categories of water load concentrations for ten chemical substance groups as defined by the Lower Saxony (Germany) Ministry of the Environment: class I (no load), class I-II (very minor), class II (minor, which is the target-setting by the Ministry)), class II-III (moderate to substantial), class III (substantial), class III-IV (strong), class IV (very strong). With the exception for AOX (µg/l), all concentrations are given in mg per liter.

	class I	class I-II	class II	class II-III	class III	class III-IV)	class IV
Total nitrogen	< 1	< 1.5	< 3	< 6	< 12	< 24	> 24
Ammonium	< 0.04	< 0.1	< 0.3	< 0.6	< 1.2	< 2.4	> 2.4
Nitrate	< 1	< 1.5	< 2.5	< 5	< 10	< 20	> 20
Nitrite	< 0.01	< 0.05	< 0.1	< 0.2	< 0.4	< 0.8	> 0.8
AOX	< 0	< 10	< 25	< 50	< 100	< 200	> 200
Phosphate	< 0.02	< 0.04	< 0.1	< 0.2	< 0.4	< 0.8	> 0.8
Sulfate	< 25	< 50	< 100	< 200	< 400	< 800	> 800
TOC	< 2	< 3	< 5	< 10	< 20	< 40	> 40
Total phosphate	< 0.05	< 0.08	< 0.15	< 0.3	< 0.6	< 1.2	> 1.2
Chloride	< 25	< 50	< 100	< 200	< 400	< 800	> 800

AOX - adsorbable organic halogen compounds
TOC - total organic carbon

Supplemental Table 2. Detailes on the names, genotypic characteristics, geographical and ecological backgroud of the *P. aeruginosa* strains.

Strain	Hexadecimal code	Haplotype	Origin	Source	Additional Info	Location	Date	Country
A 5803	F429	I	Airway	(Wiehlmann *et al.*, 2007)	unk	Heidelberg	1992	Germany
245	1BAE	E	Animal	(Khan *et al.*, 2007)	Dolphin	Japan	May-03	Japan
DSM 939	049A	S13	animal water	(Wiehlmann *et al.*, 2007)	bottle	unk	bef. 1969	unk
892	3C52	K	Burn wound	(Wiehlmann *et al.*, 2007)	Int.care unit	Hannover	1990	Germany
PA14	D421		burnwound	(Wiehlmann *et al.*, 2007)	unk	Boston	unk	USA
H2	241A	S20	catheter	(Wiehlmann *et al.*, 2007)	unk	unk	unk	unk
4	F469	D	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
11	149A	E33	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
12	239A	L	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
14	6C1A	S	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
21	EC29	J1	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
36	E479	E46	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1998	Germany
54	F46A	D1	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1999	Germany
56	2C22	E27	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1999	Germany
69	E022	E17	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	unk	Germany
70	2C2A	E71	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1999	Germany
125	6C2A	P	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jan-01	Germany
126	F429	I	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jan-01	Germany
128	EA0A	A3	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Feb-01	Germany
247	C40A	C	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jun-02	Germany
252	C40A	C	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Dec-03	Germany
265	0812	V	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Nov-03	Germany
278	6C22	Y	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jun-03	Germany
321	3C2A	U	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	unk	Germany
323	E429	B	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jun-03	Germany
345	4C12	T	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Feb-03	Germany
362	EA0A	A3	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jul-08	Germany
404	0812	V	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Sep-04	Germany
424	2F82	S7	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Sep-04	Germany
433	2C22	E27	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	unk	Germany
483	0BA2	E31	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	unk	Germany
63741	3C52	K	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jan-02	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
172-5645	741E	E9	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Mar-02	Germany
188-5064	F421	A2	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Jul-02	Germany
195-5050	D421	A	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Nov-02	Germany
212-5826	D421	A	CF	(Wiehlmann *et al.*, 2007)	Sputum	Kopenhagen	1992	Denmark
2733/92	3C2A	U	CF	(Wiehlmann *et al.*, 2007)	Sputum	Kopenhagen	1992	Denmark
2813 A/92	4012	O	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	Sep-05	Germany
491-5530	EC2A	J	CF	(Wiehlmann *et al.*, 2007)	Sputum	Hannover	1983	Germany
BST 1	E469	S52	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1985	Germany
CHA	EC2A	J	CF	(Wiehlmann *et al.*, 2007)	throat swab	Grenoble	unk	France
DM	E84A	S1	CF	(Wiehlmann *et al.*, 2007)	Sputum	Hamburg	1984	Germany
G7	6D92	H	CF	(Wiehlmann *et al.*, 2007)	Sputum	Stade	1986	Germany
K9	1BAE	E	CF	(Wiehlmann *et al.*, 2007)	Sputum	Husum	1985	Germany
KB1	059A	S16	CF	(Wiehlmann *et al.*, 2007)	throat swab	Bremen	1987	Germany
LES400	4C12	T	CF	(Wiehlmann *et al.*, 2007)	throat swab	Liverpool	unk	Great Britain
MF 6	AC9A	S10	CF	(Wiehlmann *et al.*, 2007)	throat swab	Bremen	1987	Germany
PD 1	E59A	S47	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1985	Germany
RN 4	D421	A	CF	(Wiehlmann *et al.*, 2007)	throat swab	Oldenburg	1986	Germany
RP 1	0C2E	G	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1985	Germany
SG1 (= C)	C40A	C	CF	(Wiehlmann *et al.*, 2007)	throat swab	Bueckeburg	1986	Germany
SS 1	6D92	H	CF	(Wiehlmann *et al.*, 2007)	throat swab	Lueneburg	1985	Germany
TBCF10839	3C52	K	CF	(Wiehlmann *et al.*, 2007)	throat swab	Hannover	1983	Germany
Va 24437	3C51	K1	CF	(Wiehlmann *et al.*, 2007)	Sputum	Halle	1992	Germany
Va 26232	EC2A	J	CF	(Wiehlmann *et al.*, 2007)	Sputum	Halle	1992	Germany
Va 27081	081E	A8	CF	(Wiehlmann *et al.*, 2007)	Sputum	Halle	1992	Germany
Va 27260	239A	L	CF	(Wiehlmann *et al.*, 2007)	Sputum	Halle	1992	Germany
ZW 102	2E12	S26	CF	(Wiehlmann *et al.*, 2007)	throat swab	Leuven	1997	Belgium
ZW 113	6E12	M	CF	(Wiehlmann *et al.*, 2007)	throat swab	Rotterdam	1997	Netherlands
ZW 117	0812	V	CF	(Wiehlmann *et al.*, 2007)	throat swab	Vienna	1997	Austria
ZW 119	F469	D	CF	(Wiehlmann *et al.*, 2007)	throat swab	Posen	1997	Poland
ZW 30	B420	B30	CF	(Wiehlmann *et al.*, 2007)	throat swab	Innsbruck	1997	Austria
ZW 31	AC2E	S48	CF	(Wiehlmann *et al.*, 2007)	throat swab	Innsbruck	1997	Austria
ZW 41	0192	S14	CF	(Wiehlmann *et al.*, 2007)	throat swab	Verona	1997	Italy
ZW 43	3C52	K	CF	(Wiehlmann *et al.*, 2007)	throat swab	Genua	1997	Italy

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
ZW 49	A5AA	S46	CF	(Wiehlmann *et al.*, 2007)	throat swab	Verona	1997	Italy
ZW 54	6C12	S42	CF	(Wiehlmann *et al.*, 2007)	throat swab	Milano	1997	Italy
ZW 64-1	279A	S22	CF	(Wiehlmann *et al.*, 2007)	throat swab	Lund	1997	Sweden
ZW 77	4012	O	CF	(Wiehlmann *et al.*, 2007)	throat swab	London	1997	UK
ZW 79	0C2E	G	CF	(Wiehlmann *et al.*, 2007)	throat swab	Galway	1997	Ireland
ZW 81	D429	Q	CF	(Wiehlmann *et al.*, 2007)	throat swab	London	1997	UK
ZW 83	6E12	M	CF	(Wiehlmann *et al.*, 2007)	throat swab	London	1997	UK
ZW 85	D421	A	CF	(Wiehlmann *et al.*, 2007)	throat swab	Aberdeen	1997	UK
ZW 88	2C1A	F	CF	(Wiehlmann *et al.*, 2007)	throat swab	London	1997	UK
ZW 92	EC22	S8	CF	(Wiehlmann *et al.*, 2007)	throat swab	Marseille	1997	France
ZW 98a	E419	S57	CF	(Wiehlmann *et al.*, 2007)	throat swab	Den Haag	1997	Netherlands
1303	CC28	E54	Coastal	(Khan *et al.*, 2007)	Kuroshio	Japan	Jun-03	Japan
963	2C1A	F	Coastal	(Khan *et al.*, 2007)	Tokyo	Japan	May-03	Japan
242	0C2E	G	Coastal	(Khan *et al.*, 2007)	Sagami	Japan	May-03	Japan
243	0C2E	G	Coastal	(Khan *et al.*, 2007)	Sagami	Japan	May-03	Japan
601	0C2E	G	Coastal	(Khan *et al.*, 2007)	Kumamoto	Japan	Dec-03	Japan
54	049A	S13	Coastal	(Khan *et al.*, 2007)	Suruga	Japan	Mar-04	Japan
60	049A	S13	Coastal	(Khan *et al.*, 2007)	Suruga	Japan	Mar-04	Japan
M70A.2	0F9E	A7	Faeces	(Vives-Flórez & Garnica, 2006)	unk	Colombia	unk	Colombia
M86A.1	0F9A	E42	Faeces	(Vives-Flórez & Garnica, 2006)	unk	Colombia	2006	Colombia
M87A.1	0F9A	E42	Faeces	(Vives-Flórez & Garnica, 2006)	unk	Colombia	2006	Colombia
M82A.3	AC9A	S10	Faeces	(Vives-Flórez & Garnica, 2006)	unk	Colombia	unk	Colombia
ATCC 33356	CD9E	S55	Faeces	(Wiehlmann *et al.*, 2007)	unk	Heidelberg	1965	Germany
1540	8E9A	E21	Freshwater	(Khan *et al.*, 2007)	Zenpukujii	Japan	Sep-03	Japan
1549	8E9A	E21	Freshwater	(Khan *et al.*, 2007)	Zenpukujii	Japan	Sep-03	Japan
1563	3C1A	E51	Freshwater	(Khan *et al.*, 2007)	Tamako	Japan	Sep-03	Japan
p6_1	4C8E	E68	Freshwater	Claudia Scotta	fountain	Majorca	Jun-08	Spain
p6_3	4C8E	E68	Freshwater	Claudia Scotta	fountain	Majorca	Jun-08	Spain
p6_2	6E0A	E69	Freshwater	Claudia Scotta	fountain	Majorca	Jun-08	Spain
p4_2	B42A	E70	Freshwater	Claudia Scotta	fountain	Majorca	Jun-08	Spain
p4_3	B42A	E70	Freshwater	Claudia Scotta	fountain	Majorca	Jun-08	Spain
ATCC 33988	6C22	Y	fuel tank	(Wiehlmann *et al.*, 2007)	Ponca City	Oklahoma	unk	USA
28	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
29	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
30	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
31	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
32	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
33	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
35	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
36	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 2	USA	1987	USA
38	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 2	USA	1987	USA
40	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
41	EA0A	A3	G. c. aquifer	(Foght *et al.*, 1996)	Aq 8	USA	1987	USA
PAK	55AA	S39	laboratory	(Wiehlmann *et al.*, 2007)	unk	unk	bef. 1960	Japan
M28A.1	4C12	T	Manure	(Vives-Flórez & Garnica, 2006)	unk	Colombia	unk	Colombia
1155	E429	B	Open Ocean	(Khan *et al.*, 2007)	Ocean	Japan	Jun-03	Japan
1187	E429	B	Open Ocean	(Khan *et al.*, 2007)	Ocean	Japan	Jun-03	Japan
1200	E429	B	Open Ocean	(Khan *et al.*, 2007)	Ocean	Japan	Jun-03	Japan
1206	E429	B	Open Ocean	(Khan *et al.*, 2007)	Ocean	Japan	Jun-03	Japan
100	EC21	E52	Open Ocean	(Khan *et al.*, 2007)	Ocean	Japan	Mar-04	Japan
DSM 1128	EC38	S32	Patient	(Wiehlmann *et al.*, 2007)	ear infect.	unk	unk	unk
206	4C23	E53	River water	(Khan *et al.*, 2007)	Arakawa	Japan	May-03	Japan
1508	3C2A	U	River water	(Khan *et al.*, 2007)	Arakawa	Japan	Nov-03	Japan
JeDePA	E661	E40	Soil	unk	unk	India	unk	India
M41A.1	478A	E41	Soil	(Vives-Flórez & Garnica, 2006)	unk	Colombia	unk	Colombia
M9A.1	2C0A	E44	Soil	(Vives-Flórez & Garnica, 2006)	Capachos	Colombia	unk	Colombia
M15A.3	EFAE	E45	Soil	(Vives-Flórez & Garnica, 2006)	St. Marta	Colombia	unk	Colombia
BTX28A	2C2A	E71	Soil	(Vives-Flórez & Garnica, 2006)	unk	Colombia	unk	Colombia
M19A.1	6C22	Y	Soil	(Vives-Flórez & Garnica, 2006)	Garzas	Colombia	unk	Colombia
ATCC 10145	46BA	S38	type strain	(Wiehlmann *et al.*, 2007)	unk	unk	1963	Czehoslovak.
ATCC 33348	2C1A	F	unk	(Wiehlmann *et al.*, 2007)	unk	unk	bef. 1980	Germany
ATCC 33364	E42A	B	unk	(Wiehlmann *et al.*, 2007)	unk	unk	1982	unk
DSM 1253	0B92	X	unk	(Wiehlmann *et al.*, 2007)	unk	Stanford	1949	USA
DSM 288	0B92	X	unk	(Wiehlmann *et al.*, 2007)	Hygiene inst.	Goettingen	1990	Germany
Gr 2057	2C92	S31	unk	(Wiehlmann *et al.*, 2007)	unk	Athens	1995	Greece
Gr 2052	2C92	S31	unk	(Wiehlmann *et al.*, 2007)	unk	Athens	1995	Greece

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
Gr 2248	6D92	H	unk	(Wiehlmann *et al.*, 2007)	unk	Athens	1995	Greece
HJ2	0C2E	G	unk	(Wiehlmann *et al.*, 2007)	Sputum	Cologne	1990	Germany
429SF6	2C9A	E49	Water	This study	W	Ströhen	Jul-08	Germany
110E8	D421	A	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B13	F661	A1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
114SE1	4C2E	A11	Water	This study	O/A	Börßum	Jul-08	Germany
055E6	F421	A2	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
012SD1	EA0A	A3	Water	This study	W	Hann. Münden	Jul-08	Germany
120D11	85AA	A4	Water	This study	O/A	Peine	Oct-07	Germany
012B5	7C2E	A5	Water	This study	W	Hann. Münden	Oct-07	Germany
702SC1	E429	B	Water	This study	UW	Langwedel	Jul-08	Germany
120SD3	B420	B30	Water	This study	O/A	Peine	Jul-08	Germany
120SF7	B021	B31	Water	This study	O/A	Peine	Jul-08	Germany
010SB1	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
110E7	F469	D	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
406E1	F46A	D1	Water	This study	W	Drakenburg	Oct-07	Germany
110B2	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
406B1	8D9A	E10	Water	This study	W	Drakenburg	Oct-07	Germany
429C2	B020	E11	Water	This study	W	Ströhen	Oct-07	Germany
120SE7	CF92	E12	Water	This study	O/A	Peine	Jul-08	Germany
429C4	4E22	E13	Water	This study	W	Ströhen	Oct-07	Germany
406C1	CC9A	E14	Water	This study	W	Drakenburg	Oct-07	Germany
120B3	AE1A	E15	Water	This study	O/A	Peine	Oct-07	Germany
429C1	FD82	E16	Water	This study	W	Ströhen	Oct-07	Germany
429E4	E022	E17	Water	This study	W	Ströhen	Oct-07	Germany
702D4	885A	E18	Water	This study	UW	Langwedel	Oct-07	Germany
702F3	FC4A	E19	Water	This study	UW	Langwedel	Oct-07	Germany
110D1	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
120SF3	6E1A	E20	Water	This study	O/A	Peine	Jul-08	Germany
924SE4	6D9A	E22	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
110SB6	8E9A	E21	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
401F1	062E	E24	Water	This study	W	Allerbruch	Oct-07	Germany
706B1	282A	E23	Water	This study	O/A	Celle	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
110F3	4822	E25	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B14	05AA	E26	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
012A2	2C22	E27	Water	This study	W	Hann.Münden	Oct-07	Germany
012D3	0C1E	E28	Water	This study	W	Hann.Münden	Oct-07	Germany
055E11	499A	E32	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
902C1	4B9A	E3	Water	This study	UW	Brake-Weser	Oct-07	Germany
055F1	E422	E30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055C1	0BA2	E31	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055E7	499A	E32	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
120SB2	149A	E33	Water	This study	O/A	Peine	Jul-08	Germany
114A2	A99A	E34	Water	This study	O/A	Börßum	Oct-07	Germany
120E2	E421	E35	Water	This study	O/A	Peine	Oct-07	Germany
120E9	E99A	E36	Water	This study	O/A	Peine	Oct-07	Germany
120D8	659A	E37	Water	This study	O/A	Peine	Oct-07	Germany
012B2	259A	E38	Water	This study	W	Hann.Münden	Oct-07	Germany
110SE2	002A	E39	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
706D1	EF9A	E4	Water	This study	O/A	Celle	Oct-07	Germany
120SC3	0F9A	E42	Water	This study	O/A	Peine	Jul-08	Germany
924SE1	2E1A	E47	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
012SF1	2C9A	E49	Water	This study	W	Hann. Münden	Jul-08	Germany
120SD1	0C92	E5	Water	This study	O/A	Peine	Jul-08	Germany
055SE4	0D9E	E55	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
110SA1	AFBA	E56	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110SB4	262A	E57	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
055SE2	EC12	E58	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SD3	F46A	D1	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
114SB4	D469	E6	Water	This study	O/A	Börßum	Jul-08	Germany
702SE1	6822	E60	Water	This study	UW	Langwedel	Jul-08	Germany
110SA3	6BAE	E63	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
702SD3	679E	E61	Water	This study	UW	Langwedel	Jul-08	Germany
110SB3	0E9A	E62	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
902SA4	0322	E65	Water	This study	UW	Brake-Weser	Jul-08	Germany
706SA2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
902SD3	B411	E66	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SE1	E82E	E67	Water	This study	UW	Brake-Weser	Jul-08	Germany
120SE2	82BA	E7	Water	This study	O/A	Peine	Jul-08	Germany
110SA2	2C2A	E71	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
902F1	E41B	E73	Water	This study	UW	Brake-Weser	Oct-07	Germany
120B5	F49B	E8	Water	This study	O/A	Peine	Oct-07	Germany
055SD1	741E	E9	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
110SF5	2C1A	F	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
429E5	0C2E	G	Water	This study	W	Ströhen	Oct-07	Germany
406A2	6D92	H	Water	This study	W	Drakenburg	Oct-07	Germany
120SF5	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SD2	EC2A	J	Water	This study	O/A	Peine	Jul-08	Germany
110SC1	EC29	J1	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
114SC1	3C52	K	Water	This study	O/A	Börßum	Jul-08	Germany
012SC1	239A	L	Water	This study	W	Hann. Münden	Jul-08	Germany
110E6	6E12	M	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
114SB3	6C2A	P	Water	This study	O/A	Börßum	Jul-08	Germany
012SE1	6C1A	S	Water	This study	W	Hann. Münden	Jul-08	Germany
429A2	E84A	S1	Water	This study	W	Ströhen	Oct-07	Germany
924SF5	0D9A	S19	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
706SC3	2C12	S27	Water	This study	O/A	Celle	Jul-08	Germany
406SF1	2C2E	S29	Water	This study	W	Drakenburg	Jul-08	Germany
702SF2	0C1A	S3	Water	This study	UW	Langwedel	Jul-08	Germany
114SA7	2FAA	S35	Water	This study	O/A	Börßum	Jul-08	Germany
702E1	6CA2	S4	Water	This study	UW	Langwedel	Oct-07	Germany
012SC3	841E	S44	Water	This study	W	Hann. Münden	Jul-08	Germany
902SB4	4992	S5	Water	This study	UW	Brake-Weser	Jul-08	Germany
924SB3	EC22	S8	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
110SD4	CD9E	S55	Water	This study	O/A	Gross Schw	Jul-08	Germany
120A1	8C2A	S6	Water	This study	O/A	Peine	Oct-07	Germany
120F4	2F82	S7	Water	This study	O/A	Peine	Oct-07	Germany
429D2	4C12	T	Water	This study	W	Ströhen	Oct-07	Germany
012E6	882A	S9	Water	This study	W	Hann.Münden	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
120D6	3C2A	U	Water	This study	O/A	Peine	Oct-07	Germany
429SC1	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
120SA2	6C22	Y	Water	This study	O/A	Peine	Jul-08	Germany
012D2	D421	A	Water	This study	W	Hann.Münden	Oct-07	Germany
012F10	D421	A	Water	This study	W	Hann.Münden	Oct-07	Germany
055F5	D421	A	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055F6	D421	A	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
120B6	D421	A	Water	This study	O/A	Peine	Oct-07	Germany
012SA1	D421	A	Water	This study	W	Hann. Münden	Jul-08	Germany
503	D421	A	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
114D1	F421	A2	Water	This study	O/A	Börßum	Oct-07	Germany
055SA1	F421	A2	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
506	F421	A2	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
012E2	EA0A	A3	Water	This study	W	Hann.Münden	Oct-07	Germany
055B1	EA0A	A3	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055E10	EA0A	A3	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055F4	EA0A	A3	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
110B6	EA0A	A3	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B9	EA0A	A3	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110D12	EA0A	A3	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110E2	EA0A	A3	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
114B1	EA0A	A3	Water	This study	O/A	Börßum	Oct-07	Germany
114D2	EA0A	A3	Water	This study	O/A	Börßum	Oct-07	Germany
114E1	EA0A	A3	Water	This study	O/A	Börßum	Oct-07	Germany
702A1	EA0A	A3	Water	This study	UW	Langwedel	Oct-07	Germany
702F1	EA0A	A3	Water	This study	UW	Langwedel	Oct-07	Germany
055E1	85AA	A4	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
706C2	EA0A	A3	Water	This study	O/A	Celle	Oct-07	Germany
012SE3	EA0A	A3	Water	This study	W	Hann. Münden	Jul-08	Germany
W15Aug30	EA0A	A3	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
012D1	85AA	A4	Water	This study	W	Hann.Münden	Oct-07	Germany
114F1	85AA	A4	Water	This study	O/A	Börßum	Oct-07	Germany
055E2	85AA	A4	Water	This study	W	Hessisch Oldend.	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
523	0F9E	A7	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
055F2	E429	B	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
406A1	E429	B	Water	This study	W	Drakenburg	Oct-07	Germany
513	E429	B	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
520	E429	B	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
055D2	B420	B30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055D3	B420	B30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055D4	B420	B30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055E3	B420	B30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055E5	B420	B30	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
110D9	B420	B30	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110E1	B420	B30	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
120C1	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120D4	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120D5	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120D9	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120D10	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120E4	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120E5	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
120F2	B420	B30	Water	This study	O/A	Peine	Oct-07	Germany
401D1	B420	B30	Water	This study	W	Allerbruch	Oct-07	Germany
429E3	B420	B30	Water	This study	W	Ströhen	Oct-07	Germany
429F1	B420	B30	Water	This study	W	Ströhen	Oct-07	Germany
702E2	B420	B30	Water	This study	UW	Langwedel	Oct-07	Germany
706F1	B420	B30	Water	This study	O/A	Celle	Oct-07	Germany
114SF4	B420	B30	Water	This study	O/A	Börßum	Jul-08	Germany
924D1	B420	B30	Water	This study	UW	Brake-Sieltief	Oct-07	Germany
114SC3	B420	B30	Water	This study	O/A	Börßum	Jul-08	Germany
114SD2	B420	B30	Water	This study	O/A	Börßum	Jul-08	Germany
114SE2	B420	B30	Water	This study	O/A	Börßum	Jul-08	Germany
114SF1	B420	B30	Water	This study	O/A	Börßum	Jul-08	Germany
902SA3	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SA1	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
902SB2	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SB3	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SC1	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SD1	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SD2	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SE3	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SF2	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SF3	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SF4	B420	B30	Water	This study	UW	Brake-Weser	Jul-08	Germany
924SA3	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SB1	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SD2	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SD5	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SF3	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SF6	B420	B30	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
702SB3	B021	B31	Water	This study	UW	Langwedel	Jul-08	Germany
702SF3	B021	B31	Water	This study	UW	Langwedel	Jul-08	Germany
010A1	C40A	C	Water	This study	O/A	Goslar	Oct-07	Germany
010C1	C40A	C	Water	This study	O/A	Goslar	Oct-07	Germany
010D1	C40A	C	Water	This study	O/A	Goslar	Oct-07	Germany
010E1	C40A	C	Water	This study	O/A	Goslar	Oct-07	Germany
012D4	C40A	C	Water	This study	W	Hann.Münden	Oct-07	Germany
012F8	C40A	C	Water	This study	W	Hann.Münden	Oct-07	Germany
012F9	C40A	C	Water	This study	W	Hann.Münden	Oct-07	Germany
010SD1	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
110D4	C40A	C	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
702D3	C40A	C	Water	This study	UW	Langwedel	Oct-07	Germany
902D1	C40A	C	Water	This study	UW	Brake-Weser	Oct-07	Germany
010SA1	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
010SB2	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
010SC1	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
010SE3	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
010SE2	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
010SE4	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
010SF2	C40A	C	Water	This study	O/A	Goslar	Jul-08	Germany
012SB1	C40A	C	Water	This study	W	Hann. Münden	Jul-08	Germany
012SC2	C40A	C	Water	This study	W	Hann. Münden	Jul-08	Germany
W5Aug28	C40A	C	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
510	C40A	C	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
114SA1	F469	D	Water	This study	O/A	Börßum	Jul-08	Germany
W15Dec14	F469	D	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
429A1	F46A	D1	Water	This study	W	Ströhen	Oct-07	Germany
429B1	F46A	D1	Water	This study	W	Ströhen	Oct-07	Germany
702D1	F46A	D1	Water	This study	UW	Langwedel	Oct-07	Germany
902A1	F46A	D1	Water	This study	UW	Brake-Weser	Oct-07	Germany
902E3	F46A	D1	Water	This study	UW	Brake-Weser	Oct-07	Germany
924E1	F46A	D1	Water	This study	UW	Brake-Sieltief	Oct-07	Germany
055SF4	F46A	D1	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
114SA2	F46A	D1	Water	This study	O/A	Börßum	Jul-08	Germany
406SF3	F46A	D1	Water	This study	W	Drakenburg	Jul-08	Germany
702SA1	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SA2	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SB1	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SB2	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SE2	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SE3	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
110A5	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
702SSE5	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
702SF4	F46A	D1	Water	This study	UW	Langwedel	Jul-08	Germany
924SF7	F46A	D1	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
508	F46A	D1	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
512	F46A	D1	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
110A2	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110A4	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110A8	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110A6	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
110B1	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B5	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B7	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B11	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B12	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110C2	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110C4	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110D2	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110D11	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110E10	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110F5	CBA3	E1	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
924SA1	CBA3	E1	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SC2	CBA3	E1	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
429C3	B020	E11	Water	This study	W	Ströhen	Oct-07	Germany
706SA1	CC9A	E14	Water	This study	O/A	Celle	Jul-08	Germany
429D1	AE1A	E15	Water	This study	W	Ströhen	Oct-07	Germany
429E1	AE1A	E15	Water	This study	W	Ströhen	Oct-07	Germany
110A3	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110B8	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110D8	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110E3	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110E4	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
120E3	6E1A	E20	Water	This study	O/A	Peine	Oct-07	Germany
110E9	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
110F2	081A	E2	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
114SA3	081A	E2	Water	This study	O/A	Börßum	Jul-08	Germany
114SA4	081A	E2	Water	This study	O/A	Börßum	Jul-08	Germany
114SD1	081A	E2	Water	This study	O/A	Börßum	Jul-08	Germany
114SF3	081A	E2	Water	This study	O/A	Börßum	Jul-08	Germany
120B2	6E1A	E20	Water	This study	O/A	Peine	Oct-07	Germany
120D7	6E1A	E20	Water	This study	O/A	Peine	Oct-07	Germany
702F2	6E1A	E20	Water	This study	UW	Langwedel	Oct-07	Germany
120F1	6E1A	E20	Water	This study	O/A	Peine	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
702D2	8E9A	E21	Water	This study	UW	Langwedel	Oct-07	Germany
110SB2	8E9A	E21	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110D3	6D9A	E22	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
429SF1	6D9A	E22	Water	This study	W	Ströhen	Jul-08	Germany
114D3	2C22	E27	Water	This study	O/A	Börßum	Oct-07	Germany
012SD2	2C22	E27	Water	This study	W	Hann. Münden	Jul-08	Germany
516	2C22	E27	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
511	0C1E	E28	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
055E9	499A	E32	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
012E1	149A	E33	Water	This study	W	Hann.Münden	Oct-07	Germany
012E4	149A	E33	Water	This study	W	Hann.Münden	Oct-07	Germany
012E5	149A	E33	Water	This study	W	Hann.Münden	Oct-07	Germany
012F7	149A	E33	Water	This study	W	Hann.Münden	Oct-07	Germany
055F3	149A	E33	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
110SF1	149A	E33	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
W15Dec3	149A	E33	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
W15Dec2	149A	E33	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
518	149A	E33	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
517	149A	E33	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
114A1	A99A	E34	Water	This study	O/A	Börßum	Oct-07	Germany
W5Aug16	259A	E38	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
519	2E1A	E47	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
012F1	002A	E39	Water	This study	W	Hann.Münden	Oct-07	Germany
012F2	002A	E39	Water	This study	W	Hann.Münden	Oct-07	Germany
M8A.1	0F9A	E42	Water	(Vives-Flórez & Garnica, 2006)	Cano Limon	Colombia	2001	Colombia
M8A.3	0F9A	E42	Water	(Vives-Flórez & Garnica, 2006)	Cano Limon	Colombia	2001	Colombia
M8A.2	0F9A	E42	Water	(Vives-Flórez & Garnica, 2006)	Cano Limon	Colombia	2001	Colombia
M8A.4	2822	E43	Water	(Vives-Flórez & Garnica, 2006)	Cano Limon	Colombia	2001	Colombia
524	E479	E46	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
499	E479	E46	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
525	2E1A	E47	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
505	7C1A	E50	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
497	C81A	E48	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
110SB5	AFBA	E56	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110SC2	AFBA	E56	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110SD2	262A	E57	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
902SC4	262A	E57	Water	This study	UW	Brake-Weser	Jul-08	Germany
902E2	D469	E6	Water	This study	UW	Brake-Weser	Oct-07	Germany
406SF2	6822	E60	Water	This study	W	Drakenburg	Jul-08	Germany
406SF4	6822	E60	Water	This study	W	Drakenburg	Jul-08	Germany
110SE3	0E9A	E62	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110SD3	0E9A	E62	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
110SF7	0E9A	E62	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
706SA3	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SB1	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SB2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SB3	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SB4	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SC2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SC4	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SD1	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SD2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SD3	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
120B1	F49B	E8	Water	This study	O/A	Peine	Oct-07	Germany
706SE1	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SE2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SE4	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SE5	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SF1	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SF2	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
706SF3	4C1A	E64	Water	This study	O/A	Celle	Jul-08	Germany
902SC2	0322	E65	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SC3	B411	E66	Water	This study	UW	Brake-Weser	Jul-08	Germany
902SE2	B411	E66	Water	This study	UW	Brake-Weser	Jul-08	Germany
120E7	F49B	E8	Water	This study	O/A	Peine	Oct-07	Germany
120D3	F49B	E8	Water	This study	O/A	Peine	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
406D2	741E	E9	Water	This study	W	Drakenburg	Oct-07	Germany
406D3	741E	E9	Water	This study	W	Drakenburg	Oct-07	Germany
406D4	741E	E9	Water	This study	W	Drakenburg	Oct-07	Germany
507	2C1A	F	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
012A1	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012A4	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012B3	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012B4	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012C1	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012E3	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
012F6	0C2E	G	Water	This study	W	Hann.Münden	Oct-07	Germany
406D1	0C2E	G	Water	This study	W	Drakenburg	Oct-07	Germany
012F5	F429	I	Water	This study	W	Hann.Münden	Oct-07	Germany
120B4	F429	I	Water	This study	O/A	Peine	Oct-07	Germany
055SB2	F429	I	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
110SB7	F429	I	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
114SE3	F429	I	Water	This study	O/A	Börßum	Jul-08	Germany
114SF2	F429	I	Water	This study	O/A	Börßum	Jul-08	Germany
120SA1	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
702SB4	EC2A	J	Water	This study	UW	Langwedel	Jul-08	Germany
120SA3	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SB1	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SB3	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SD4	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SD6	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SE5	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
120SE8	F429	I	Water	This study	O/A	Peine	Jul-08	Germany
W15Apr4	F429	I	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
PSWtB	F42A	I1	Water	(Behar *et al.*, 2008)	unk	Israel	unk	Israel
110SF4	EC2A	J	Water	This study	O/A	Groß Schwülp.	Jul-08	Germany
120SF6	EC2A	J	Water	This study	O/A	Peine	Jul-08	Germany
012F4	EC29	J1	Water	This study	W	Hann.Münden	Oct-07	Germany
514	EC2A	J	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
702F4	EC29	J1	Water	This study	UW	Langwedel	Oct-07	Germany
706B2	EC29	J1	Water	This study	O/A	Celle	Oct-07	Germany
010SE1	EC29	J1	Water	This study	O/A	Goslar	Jul-08	Germany
012SF2	EC29	J1	Water	This study	W	Hann. Münden	Jul-08	Germany
110C1	239A	L	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
429E2	239A	L	Water	This study	W	Ströhen	Oct-07	Germany
902E1	239A	L	Water	This study	UW	Brake-Weser	Oct-07	Germany
522	239A	L	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
W15Aug23	6C2A	P	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
504	7D9A	R	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
W11Aug25	6C1A	S	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
521	6C1A	S	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
496	6C1A	S	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
515	6B92	S12	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
702SF1	0D9A	S19	Water	This study	UW	Langwedel	Jul-08	Germany
924SC3	0D9A	S19	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SD1	0D9A	S19	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SD4	0D9A	S19	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
509	3C2A	U	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
924SE2	0D9A	S19	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
702E3	0C1A	S3	Water	This study	UW	Langwedel	Oct-07	Germany
012SD3	2FAA	S35	Water	This study	W	Hann. Münden	Jul-08	Germany
114SB1	2FAA	S35	Water	This study	O/A	Börßum	Jul-08	Germany
114SC2	2FAA	S35	Water	This study	O/A	Börßum	Jul-08	Germany
110D5	4992	S5	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
706C4	4992	S5	Water	This study	O/A	Celle	Oct-07	Germany
W15Dec4	2F82	S7	Water	(Pirnay *et al.*, 2005)	Wlw	Brussels	2001	Belgium
120E8	EC22	S8	Water	This study	O/A	Peine	Oct-07	Germany
902SB1	EC22	S8	Water	This study	UW	Brake-Weser	Jul-08	Germany
498	4C12	T	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
500	3C2A	U	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
055D5	0812	V	Water	This study	W	Hessisch Oldend.	Oct-07	Germany
055D1	0812	V	Water	This study	W	Hessisch Oldend.	Oct-07	Germany

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
110A7	0812	V	Water	This study	O/A	Groß Schwülp.	Oct-07	Germany
120F3	0812	V	Water	This study	O/A	Peine	Oct-07	Germany
055SA2	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SA3	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SB1	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SB3	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SB4	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SC1	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SC2	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SC3	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SD4	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SE3	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SE5	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SF2	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
055SF3	0812	V	Water	This study	W	Hess. Oldendorf	Jul-08	Germany
429SF2	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SF4	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
501	0B92	X	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan
429SF5	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SA1	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SD1	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SD2	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SE1	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SE2	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SE3	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SE4	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SE5	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
429SF3	0812	V	Water	This study	W	Ströhen	Jul-08	Germany
924SA2	0812	V	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SB2	0812	V	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
924SD3	0812	V	Water	This study	UW	Brake-Sieltief	Jul-08	Germany
A 5670	7D9A	R	Wound	(Wiehlmann *et al.*, 2007)	unk	Heidelberg	1992	Germany
502	0B92	X	Water	Khan (Mutharia *et al.*, 1982)	unk	Japan	bef. 1974	Japan

Strain	Hexadecimal code	Origin	Source	Additional Info	Location	Date	Country	Strain
AL 5846	D429	Q	Wound	(Wiehlmann *et al.*, 2007)	unk	Heidelberg	1992	Germany
ATCC 15691	7D9A	R	Wound	(Wiehlmann *et al.*, 2007)	unk	Melbourne	1952	Australia
PAO	0002	W	Wound	(Wiehlmann *et al.*, 2007)	DSM1707	unk	unk	unk

Unk – unknown

CF – cystic fibrosis

G. c. aquifer - Gasoline-contaminated aquifer

Supplemental Table 3. Overview of the characteristics and test results of *P. aeruginosa*. Susceptibility - susceptebility phenpotype to lytic phage attack

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin production	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
4	n	n	n	m	s	0	sl	sl	0	0	sl	0	N/A	4296484.426
11	n	n	n	m	s	cl	cl	sl	0	cl	sl	sl	1	7949162.619
12	n	n	n	m	n	0	sl	sl	0	0	0	0	P/V	13574296.6
14	nm	n	n	m	n	cl	cl	sl	0	cl	sl	sl	6	5725180.53
21	n	n	n	n.a.	n	0	0	sl	0	cl	cl	cl	11	12081273.42
36	n	n	n	m	n	cl	cl	cl	0	cl	sl	sl	6	12775769.86
54	m	m	m	n.a.	n	0	0	0	0	0	0	0	P/V	9964013.056
56	m	n	m	m	n	cl	cl	sl	0	cl	cl	cl	P/V	11223502.07
69	nm	n	n	n	n	0	cl	cl	sl	cl	cl	cl	N/A	9309557.77
70	n	n	n	m	s	cl	cl	cl	cl	cl	sl	sl	1	6131314.704
125	m	m	m	m	n	cl	cl	sl	0	0	0	0	5	7011026.389
126	n	n	n	m	n	0	0	cl	0	cl	cl	cl	11	11931688.73
128	n	n	n	m	i	sl	cl	0	0	cl	cl	sl	1	5266131.1
247	n	n	n	m	n	cl	cl	cl	0	cl	sl	sl	1	15575713.7
252	nm	m	n	m	n	cl	cl	0	sl	cl	cl	sl	P/V	13327767.79
265	nm	m	n	m	n	sl	cl	cl	0	0	cl	0	16	7610145.731
278	m	m	m	n.a.	n	cl	cl	cl	cl	cl	cl	sl	P/V	15835401.18
321	n	n	n	m	n	cl	cl	sl	0	sl	sl	0	6	11241853.85
323	n	n	n	m	n	0	0	cl	0	sl	cl	0	11	7767239.503
345	n	n	n	m	s	sl	cl	cl	sl	sl	sl	sl	6	9109711.681
362	nm	m	m	m	s	cl	cl	sl	0	0	0	0	5	8877349.326
404	m	n	n	n	n	0	0	0	0	0	0	0	P/V	14476458.44
424	n	n	n	n	n	cl	cl	sl	cl	0	0	0	16	13004330.25
433	m	m	m	m	n	sl	cl	sl	0	sl	sl	0	P/V	12272312.16
483	m	m	m	n.a.	n	0	sl	0	sl	cl	cl	cl	N/A	14873475.02
892	n	m	m	m	i	0	cl	0	sl	sl	sl	0	4	6128064.066
63741	n	n	n	n	n	0	0	0	0	0	0	0	P/V	9304042.553
172-5645	nm	m	m	n.a.	n	cl	cl	cl	0	cl	cl	sl	3	11285564.56
188-5064	n	n	m	n	n	cl	cl	cl	sl	cl	cl	sl	6	10139095.01
195-5050	n	n	n	n	i	sl	cl	sl	cl	0	0	sl	P/V	10146025.1

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin productio	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
212-5826	nm	m	m	n.a.	s	sl	0	cl	0	0	0	0	P/V	32879838.3
2733/92	n	m	m	n.a.	s	0	0	sl	cl	0	0	0	P/V	5223955.796
2813 A/92	m	m	m	m	i	0	0	0	0	sl	sl	0	P/V	5103674.269
491-5530	n	n	n	m	n	cl	cl	cl	cl	sl	0	0	6	6815369.572
A 5670	n	m	m	n.a.	n	cl	cl	sl	0	sl	sl	0	P/V	10998623.1
A 5803	n	n	n	m	s	0	cl	0	sl	0	sl	sl	P/V	15490454.1
AL 5846	n	n	m	m	s	0	0	0	0	0	0	0	P/V	8058366.286
ATCC 10145	n	n	m	m	n	0	0	cl	sl	cl	cl	sl	P/V	6862799.252
ATCC 15691	m	m	m	n.a.	n	cl	cl	cl	cl	cl	cl	0	P/V	12798577.68
ATCC 33348	n	n	n	n	n	cl	cl	cl	sl	sl	cl	0	1	7660683.772
ATCC 33356	n	n	m	m	n	0	0	sl	sl	sl	sl	sl	10	5163956.321
ATCC 33364	m	m	n	m	s	sl	0	sl	0	0	0	0	N/A	8385265.901
ATCC 33988	n	m	m	n.a.	n	0	cl	cl	0	0	0	0	11	7988723.58
BST 1	n	m	n	n	n	cl	cl	cl	sl	cl	cl	sl	1	7004430.261
CHA	n	n	m	m	n	cl	cl	cl	0	sl	sl	0	6	10684477.05
DM	m	m	n	n.a.	n	0	0	0	0	sl	sl	0	P/V	61874774.37
DSM 1128	n	n	h	m	s	0	0	0	0	cl	0	0	P/V	13340709.08
DSM 1253	m	m	m	m	n	cl	cl	sl	0	0	0	0	P/V	33489221.35
DSM 288	m	m	m	n.a.	n	cl	cl	cl	0	0	0	0	P/V	16716887.38
DSM 939	n	m	m	m	n	cl	cl	cl	sl	sl	sl	0	P/V	5762733.803
G7	n	n	m	n.a.	n	cl	cl	sl	sl	cl	cl	sl	P/V	8052183.157
Gr 2052	h	h	m	n	n	cl	cl	cl	cl	cl	cl	sl	3	10666548.15
Gr 2057	n	n	m	n	n	cl	cl	cl	cl	cl	cl	sl	3	10431954.81
Gr 2248	h	m	m	n.a.	n	0	0	0	0	0	0	0	5	9096307.984
H2	m	n	m	m	s	0	0	cl	0	sl	sl	0	4	9521392.689
HJ2	nm	m	m	m	n	cl	0	cl	0	cl	cl	0	P/V	8363971.749
K9	nm	n	m	n.a.	n	sl	0	sl	0	cl	cl	sl	P/V	4456757.698
KB1	n	m	n	n	n	cl	cl	cl	cl	0	0	0	5	6878501.121
LES400	nm	m	m	n.a.	n	0	0	sl	sl	sl	sl	0	P/V	10237001.47
MF 6	n	n	m	n	i	cl	cl	sl	cl	sl	sl	0	3	5548460.411
PA14	n	m	n	n	n	0	0	0	sl	0	0	0	10	7203300.171
PAK	n	m	n	m	n	sl	sl	0	sl	sl	sl	0	6	4956002.409
PAO	n	n	n	n	n	cl	cl	sl	cl	0	0	0	16	7522665.023

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin productio	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
PD 1	n	n	n	n	i	cl	cl	cl	0	0	0	0	9	4873289.963
RN 4	n	m	n	n	n	sl	sl	sl	cl	0	0	0	P/V	7500890.686
RP 1	n	n	n	n	n	cl	0	cl	0	cl	cl	0	P/V	11845263.24
SG1 (= C)	m	m	m	m	s	0	0	cl	sl	0	sl	sl	10	8789967.423
SS 1	n	m	m	n	n	cl	cl	sl	0	sl	sl	sl	3	12839788.66
TBCF10839	n	n	n	n	n	0	0	0	sl	0	0	0	P/V	14790232.98
Va 24437	n	n	n	m	i	0	0	0	0	0	0	0	4	10058987.03
Va 26232	h	m	m	n.a.	n	0	0	0	0	0	0	0	6	11780564.67
Va 27081	m	n	n	m	n	cl	cl	cl	cl	cl	cl	cl	10	9033336.498
Va 27260	h	m	n	n.a.	n	sl	sl	sl	0	sl	sl	0	5	11424110.92
ZW 102	n	n	n	n	i	0	0	0	0	0	0	0	10	5927929.47
ZW 113	n	m	m	n.a.	n	0	0	sl	0	0	0	0	P/V	10666494.8
ZW 117	n	m	n	n	n	sl	cl	cl	0	0	0	0	5	5749278.728
ZW 119	n	m	n	m	n	0	0	cl	0	cl	cl	sl	11	6244612.792
ZW 30	n	h	n	n	n	cl	cl	sl	cl	cl	cl	cl	3	7215073.608
ZW 31	n	n	n	m	n	cl	cl	sl	sl	cl	sl	sl	6	4761039.193
ZW 41	nm	m	m	n.a.	n	cl	cl	cl	0	0	0	0	P/V	7934053.573
ZW 43	m	n	n	n.a.	n	0	0	cl	0	cl	sl	0	10	4627478.462
ZW 49	nm	n	m	m	n	0	cl	sl	cl	cl	cl	sl	P/V	44011808.71
ZW 54	n	n	n	n	i	0	0	0	0	0	0	0	P/V	8253139.799
ZW 64-1	n	m	m	n.a.	n	0	0	0	sl	0	0	0	1	25092342.08
ZW 77	m	n	n	m	n	0	0	0	0	0	0	0	P/V	8316581.59
ZW 79	nm	m	m	m	n	0	cl	cl	sl	cl	cl	sl	P/V	25909720.55
ZW 81	nm	m	n	n.a.	n	0	0	sl	0	cl	cl	sl	P/V	42066883.55
ZW 83	n	m	m	m	n	0	cl	0	0	cl	cl	sl	3	14438270.85
ZW 85	n	m	n	n	i	0	0	cl	sl	cl	cl	cl	10	24980792.81
ZW 88	m	h	m	n.a.	n	0	0	0	0	0	0	0	7	27388081.27
ZW 92	n	n	n	m	n	0	cl	cl	cl	cl	cl	cl	P/V	8601681.354
ZW 98a	nm	m	m	n.a.	n	0	sl	0	0	cl	cl	0	P/V	13357752.7
110E8	m	n	n	n	n	0	0	0	cl	cl	cl	cl	10	22142222.69
406E1	n	n	n	n	n	0	sl	sl	cl	cl	cl	cl	3	15870434.96
702E1	n	n	n	n	n	0	sl	sl	cl	sl	sl	0	1	12311177.55
120E2	nm	m	n	n	n	0	0	0	0	0	0	0	7,8	6223119.057

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin productio	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
429E4	n	n	n	n.a.	n	0	0	0	cl	cl	cl	cl	3	4042693.547
012E6	n	n	n	n.a.	s	cl	cl	sl	sl	0	0	0	2	12696831.71
429E5	nm	m	m	n.a.	n	cl	cl	sl	cl	0	0	0	6	5450109.332
055E6	nm	m	m	m	n	sl	0	cl	0	0	0	0	7,8	11048671.07
110E6	n	m	n	n	n	0	0	0	0	cl	sl	cl	9	13230473.9
055E7	n	n	n	n	n	sl	cl	cl	cl	sl	sl	sl	6	7667534.814
110E7	n	m	n	n.a.	s	0	0	sl	0	0	0	0	11	8350686.617
120E9	n	m	n	n	n	0	sl	sl	0	cl	sl	cl	11	10310843.33
055E11	n	n	n	n	n	sl	cl	cl	cl	sl	sl	sl	6	8631830.285
010SB1	n	n	n	n	n	sl	sl	sl	sl	sl	sl	sl	1	9744937.81
012A2	nm	n	n	n	n	sl	sl	sl	sl	sl	sl	sl	3	7077114.141
012B2	n	n	n	n	n	cl	cl	cl	0	0	sl	0	6	12395279.17
012B5	n	m	n	m	s	0	0	0	0	0	0	0	6	17393356.1
012D3	n	n	n	n	n	cl	cl	cl	sl	0	0	0	9	11749969.14
012SC1	n	n	n	n	n	cl	cl	cl	cl	0	0	0	16	9749710.024
012SC3	n	n	n	n	i	cl	cl	cl	0	0	0	0	6	13225482.35
012SD1	n	n	n	n	n	cl	cl	cl	0	0	0	0	1	12464941.64
012SE1	m	m	n	m	n	cl	cl	cl	cl	cl	0	0	6	12025398.01
012SF1	n	n	n	n	n	cl	cl	sl	sl	sl	sl	sl	6	11058110.77
055C1	n	m	m	m	i	cl	cl	cl	0	0	0	0	5	7548527.854
055F1	nm	m	n	n.a.	s	0	0	0	0	0	0	0	4	11076144.4
055SD1	n	n	n	n	n	cl	cl	cl	0	sl	sl	0	1	8751032.876
055SD3	n	n	n	n	n	0	0	0	cl	0	0	0	13	9355135.871
055SE2	n	n	n	n	n	cl	cl	cl	sl	sl	cl	sl	1	13012447.95
055SE4	n	n	n	n	n	cl	cl	cl	cl	sl	sl	0	6	8728604.459
110B13	m	m	n	m	n	0	sl	sl	sl	sl	sl	sl	1	6529599.92
110B14	nm	m	m	n.a.	n	0	cl	0	cl	sl	sl	sl	6	14230957.18
110B2	nm	m	n	n.a.	n	0	0	0	0	0	0	0	4	12894822.27
110D1	n	n	n	n	n	0	0	0	0	0	sl	0	8	9062133.795
110F3	n	n	n	m	s	cl	cl	cl	0	0	0	0	9	7211494.941
110SA1	n	n	n	n	n	cl	cl	cl	cl	sl	sl	sl	6	15418390.7
110SA2	n	n	n	n	n	cl	cl	cl	sl	0	0	0	9	9040304.128
110SA3	n	n	n	n	n	cl	cl	sl	sl	sl	sl	sl	6	10441930.02

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin productio	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
110SB3	n	n	n	n	n	cl	cl	cl	0	sl	sl	sl	6	12083994.48
110SB4	n	n	n	n	n	0	sl	sl	sl	sl	0	0	1	10112099.89
110SB6	n	n	n	m	s	0	sl	0	0	0	0	0	6	6558733.149
110SC1	n	m	m	m	n	0	0	sl	0	sl	sl	sl	7	8280142.823
110SD4	n	n	n	n	n	0	0	cl	0	sl	sl	0	7	11443259.06
110SE2	n	n	n	n	n	cl	cl	cl	sl	sl	sl	sl	6	11417655.09
110SF5	n	n	n	n	n	cl	cl	0	cl	0	sl	0	1	9545111.481
114A2	n	n	m	m	n	cl	cl	cl	cl	0	0	0	9	11167389.53
114SA7	n	n	n	n	n	sl	sl	sl	sl	sl	sl	sl	6	12504704.16
114SB3	n	h	n	i	n	0	0	0	0	0	0	0	N/A	6092621.238
114SB4	n	n	n	m	s	sl	sl	sl	0	0	0	0	1	8244620.952
114SC1	n	n	n	n	n	0	0	0	0	0	0	0	4	14408483.54
114SE1	n	n	n	n	n	0	0	0	0	0	0	0	6	15468837.85
120A1	nm	n	n	m	i	0	0	0	cl	sl	sl	0	6	9784631.168
120B3	nm	n	n	n	n	0	0	sl	cl	0	0	0	4	8669938.833
120B5	n	n	n	n	n	cl	cl	sl	0	0	0	0	N/A	9203909.057
120D11	nm	n	n	n	n	cl	cl	cl	0	sl	sl	sl	6	9541896.412
120D6	n	n	n	m	i	cl	cl	0	0	sl	sl	0	6	11062201.24
120D8	nm	n	m	m	i	cl	cl	sl	sl	sl	sl	sl	6	11971619.21
120F4	n	n	n	n	n	cl	cl	cl	cl	0	0	0	16	9465665.482
120SA2	n	n	n	n	n	cl	cl	cl	sl	0	sl	0	1	8624921.937
120SB2	n	n	n	n	n	0	sl	sl	sl	sl	sl	sl	1	11376503.39
120SC3	n	n	n	n	n	0	sl	0	sl	sl	sl	sl	6	11303679.21
120SD1	n	n	n	n	n	0	0	0	sl	0	0	0	4	12291952.34
120SD2	n	n	n	n	n	cl	cl	0	0	sl	sl	0	6	10584657.75
120SD3	n	n	n	n	n	0	sl	0	0	sl	sl	sl	3	10481813.08
120SE2	n	n	n	n	n	cl	cl	cl	0	sl	sl	0	6	15052281.87
120SE7	n	n	n	n	n	cl	cl	cl	0	sl	sl	sl	6	12916653.91
120SF3	m	n	n	n	n	cl	cl	sl	sl	sl	sl	sl	6	9357455.78
120SF5	n	n	n	m	n	cl	cl	cl	0	0	0	0	1	11791946.84
120SF7	n	n	n	m	i	cl	cl	cl	0	sl	sl	sl	N/A	6727863.176
401F1	n	n	n	n	n	0	sl	sl	0	0	0	0	13	8934876.734
406A2	n	n	n	n	n	cl	cl	cl	sl	sl	cl	cl	3	11362066.21

Strain (see Supplemental Table 2)	Motility			Proteolytic activity	Pyocyanin productio	Susceptibility							Serotype	Biovolume [µm3]
	Swimming	Swarming	Twitching			4	5	3	24	25	28	26		
406B1	n	n	m	m	n	cl	cl	0	sl	sl	sl	0	6	13417307.56
406C1	n	n	m	n	n	0	sl	0	0	sl	sl	0	1	7941823.488
406SF1	n	n	n	n	n	cl	cl	cl	cl	0	0	0	N/A	6606565.948
429A2	n	n	n	n	n	0	0	sl	0	cl	sl	sl	11	13341284.12
429C1	n	n	n	n	n	cl	cl	0	sl	0	0	0	3	11217678.58
429C2	n	n	n	m	i	0	0	0	0	cl	sl	0	16	9384118.919
429C4	n	m	n	n	n	cl	cl	cl	0	cl	0	cl	6	11374920.56
429D2	n	n	n	n	n	0	sl	0	sl	0	sl	0	7,8	4930902.342
429SC1	n	n	n	n	n	sl	cl	sl	0	sl	sl	0	5	5784304.604
702D4	n	n	n	n	n	0	0	0	0	0	0	0	6	10565151.94
702F3	n	n	n	n	n	0	0	sl	sl	0	0	0	10	9896405.705
702SC1	nm	n	m	m	n	0	0	0	0	0	0	0	7	5270442.889
702SD3	n	n	n	n	n	0	0	sl	0	0	0	0	6	14455355.97
702SE1	n	n	n	n	n	cl	cl	cl	sl	sl	sl	sl	6	12625959.89
702SF2	n	n	n	n	n	0	sl	sl	0	sl	sl	0	6	11999649.79
706B1	n	m	n	n	n	cl	cl	cl	cl	sl	sl	sl	6	10102227.44
706D1	n	n	m	n	n	cl	cl	cl	cl	sl	sl	0	6	11106864.42
706SA2	n	n	n	n	n	sl	cl	0	0	0	sl	0	6	7406805.946
706SC3	nm	m	m	m	n	0	0	0	0	0	0	0	4	18341354.91
902C1	m	m	n	m	n	0	0	sl	0	0	0	0	6	9125826.548
902F1	n	n	n	m	i	0	0	0	0	0	0	0	N/A	4942543.382
902SA4	m	n	n	n	n	cl	cl	cl	0	0	0	0	16	11355171.7
902SB4	n	n	n	n	n	0	sl	sl	0	0	0	0	N/A	11350097.14
902SD3	n	n	n	n	n	0	sl	0	cl	sl	cl	cl	3	11089395.94
902SE1	n	n	n	n	n	0	sl	sl	sl	sl	sl	sl	1	9609800.174
924SB3	n	n	n	n	n	sl	cl	sl	sl	sl	sl	sl	1	11034579.7
924SE1	n	n	n	m	n	0	0	0	0	0	0	0	4	13217736.14
924SE4	n	n	n	n	n	0	sl	sl	sl	0	0	0	9	11623413.59
924SF5	n	n	n	n	n	cl	cl	cl	sl	sl	0	0	9	8827400.154

Motility. Swimming: nm – “non motile”; m – “minimal”; n – “normal”; h – “hyper”. Swarming: m - “minimal”; n - “normal”; h - “hyper”. Twitching: m - ”minimal”; n - “normal”; h - “hyper”. Proteolytic activity: n.a. - ”no activity”; m - “minimal”; n - “normal”; i - “increased” activity. Pyocyanin production: n - “normal”, i - “increased”, s - “strong” production. Phages susceptibility: 0 - no lysis; sl - semilysis; cl - clear lysis.

7. 1. References of the Appendix:

Behar, A., Yuval, B. & Jurkevitch, E. (2008). Gut bacterial communities in the Mediterranean fruit fly (*Ceratitis capitata*) and their impact on host longevity. *J Insect Physiol* **54**, 1377-1383.

Foght, J. M., Westlake, D. W., Johnson, W. M. & Ridgway, H. F. (1996). Environmental gasoline-utilizing isolates and clinical isolates of *Pseudomonas aeruginosa* are taxonomically indistinguishable by chemotaxonomic and molecular techniques. *Microbiology* **142 (Pt 9)**, 2333-2340.

Khan, N. H., Ishii, Y., Kimata-Kino, N., Esaki, H., Nishino, T., Nishimura, M. & Kogure, K. (2007). Isolation of *Pseudomonas aeruginosa* from Open Ocean and Comparison with Freshwater, Clinical, and Animal Isolates. *Microb Ecol*.

Mutharia, L. M., Nicas, T. I. & Hancock, R. E. (1982). Outer membrane proteins of *Pseudomonas aeruginosa* serotype strains. *J Infect Dis* **146**, 770-779.

Pirnay, J. P., Matthijs, S., Colak, H. & other authors (2005). Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. *Environ Microbiol* **7**, 969-980.

Vives-Flórez, M. & Garnica, D. (2006). Comparison of virulence between clinical and environmental *Pseudomonas aeruginosa* isolates. *Int Microbiol* **9**, 247-252.

Wiehlmann, L., Wagner, G., Cramer, N. & other authors (2007). Population structure of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **104**, 8101-8106.

Acknowledgements

I wish to express my sincere gratitude to my advisor, Dr. Johannes Sikorski, for introducing me to the world of evolution and speciation and for giving me the opportunity to conduct my PhD.

I would like to thank Prof. Dr. Kenneth N. Timmis for giving me a chance to execute my PhD research in his group, for his scientific advices and comments on my work.

I also wish to thank Prof. Dr. Susanne Häußler for her invaluable suggestions during my committee meeting and for the help in conducting some actual lab experiments. My thanks go to Mathias Mueskeln who, in spite of tremendous work loads still found time for me.

My big special thanks go to Dr. Lutz Wiehlmann from the MHH for his readiness to answer any of my numerous questions and for kindly teaching me how to deal with the microarray techniques.

I owe a tremendous debt of gratitude to all the fellow graduate students and post-docs with whom I have shared my career in science. My lab mates, ex- and present ones: Olga Golyshina, Tanja Chernikova, Sagrario Arias Rivas, Montri Yasawong, Simrita Cheema and Agnes Waliczek. You guys really made the lab a wonderful place to work and learn in! I am extremely grateful to all members of the UMW group, especially Gabriella Molinari, Monica Bassas Galia, Thorben Dammeyer, Daniela Näther and Olivier Duplex Ngatchou Djao. I also thank Marlon Kazmierchak for his help.

A huge hug for my parents – Mama, Papa – I deeply thank you for all your love, unlimited support and eagerness to listen to my occasional whining and despair of completing a PhD: Я вас люблю! I am most grateful to my big sis: Zoe, without your support, straightforward thinking, criticism, and love I would not be able to get to this point. I cannot find the right words to express all my feelings… I just want you to know that I will be always there for you, the same as you were there for me during the complicated and not easy time of my PhD.